The Dissertation Committee for Natalie Sara Kretsch Certifies that this is the approved version of the following dissertation:

Pubertal Development and Adolescent Risk-Taking: Understanding Individual Differences

Committee:

Kathryn Paige Harden, Supervisor

Shannon E. Cavanagh

Kim Fromme

Robert A. Josephs

David S. Yeager

Pubertal Development And Adolescent Risk-Taking: Understanding Individual Differences

by

Natalie Sara Kretsch, B.A., M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2016

Acknowledgements

I thank my advisor and mentor, Paige Harden, and the members of my dissertation committee, whose support greatly improved this project. Study 1 used data from the Texas Twin Project, supported by grant 1-R21- AA020588 to Drs. Paige Harden and Elliot Tucker-Drob, and grant 5-R24-HD042849 to the Population Research Center at the University of Texas at Austin. Many thanks to the staff, research assistants, and graduate students in the Harden and Tucker-Drob labs for their tireless efforts in data collection. Special thanks to Daniel Briley, Marie Carlson, Amanda Cheung, Patrick Quinn, Frank Mann, Laura Engelhardt, Megan Patterson for valuable feedback and support. I have also received support from the Waggoner Center for Alcohol and Addiction Studies at the University of Texas at Austin. Many thanks to Drs. Jane Mendle, and Jessica D. Cance for their contributions as co-authors on Study 2, in press in the *Journal of Research on Adolescence*, and Study 3, in press in the *Journal of Youth and Adolescence*. Dr. Laurence Steinberg at Temple University provided the computerized tasks used in Study 1. This research used data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill and funded by Grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Information on how to obtain the Add Health data files is available on the Add Health website (www.cpc.unc.edu/addhealth). No direct support was received from Grant P01-HD31921 for this project.

Finally, thank you to my parents and sister for their unconditional positive regard.

Pubertal development and adolescent risk-taking: Understanding individual differences

Natalie Sara Kretsch, Ph.D.

The University of Texas at Austin, 2016

Supervisor: Kathryn Paige Harden

This set of projects focused on individual differences—specifically, how variation in the timing, context, and perception of this universal milestone might contribute to individual differences in risky behavior. Study 1 looked at testosterone as a potential endophenotype for substance use in adolescence. Combining self-report, hormonal, and behavioral measures, this study used a twin design to test the hypothesis that testosterone mediated genetic risk for substance use via its effect on reward seeking. The primary hypothesis was not supported, as there were no phenotypic associations between testosterone, reward seeking, and initiation of substance use. Study 2 focused on girls' perceived pubertal timing in the context of their peer group, testing whether peer delinquency moderated the association between pubertal timing and delinquency. A twin comparison design was used to control for unmeasured between-family differences (family-level genetic and environmental selection effects) that would affect both peer and individual delinquency. Pubertal timing moderated the quasi-causal association between peer and individual delinquency: girls with earlier perceived pubertal timing were more similar to their nominated friends in delinquency. This interaction was only found for relative pubertal timing (asking girls to compare their development to their peers) and not

for age-standardized ratings of body changes or for age at menarche. Study 3 examined whether pubertal timing reported by one's friends and schoolmates related to perceived pubertal timing. Results showed gender differences: boys appeared similar to their peers in perceived body changes and girls appeared similar to peers in perceived relative pubertal timing. Collectively, these 3 studies highlight complexity inherent in studying sources of individual differences at a stage when numerous changes—biological, psychological, social—are underway. Understanding the extent to which these concurrent changes may or may not interact is an important step toward identifying factors that make some children prone to risk behavior.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Adolescence has been described as a health paradox. The second decade of life is one of the healthiest periods in life in terms of physical health, yet the rates of morbidity and mortality increase by 200% during this time (Forbes & Dahl, 2010). Morbidity and mortality during this period are primarily due to health risk behavior, including substance use, reckless driving, unprotected sex, and delinquency (Resnick et al., 1997; CDC 2009). The high rate of risky behavior among adolescents compared to children and adults is not the result of lack of knowledge about potential consequences of risky behavior or deficiencies in information processing; adolescents engage in these behaviors despite knowledge of their potential negative consequences (Beyth-Marom et al., 1993; Reyna & Farley, 2006; Steinberg, 2008). Moreover, while *mean* levels of risky behavior increase during this time, peak in young adulthood, and decline thereafter, there are marked individual differences in propensity for risk-taking, and these behaviors tend to cluster together within individuals (Caspi & Moffit, 1993; Jessor & Jessor, 1977). Thus, understanding why adolescents are risk-prone on average and why certain adolescents are more at risk than others is a challenge of major public health significance.

The increasing propensity to engage in risky behavior during adolescence may be linked to the biological changes that occur at puberty. Puberty is a universal developmental milestone, yet adolescents vary in when and how they experience this transition. The aim of this dissertation was to explore the association between adolescent risk-taking and pubertal development. Individual differences in the nature, context, and timing of puberty were expected to relate to individual differences in substance use and risk-taking. This project includes three studies, each employing a unique methodology and investigating unique outstanding questions regarding puberty and risk-taking. This

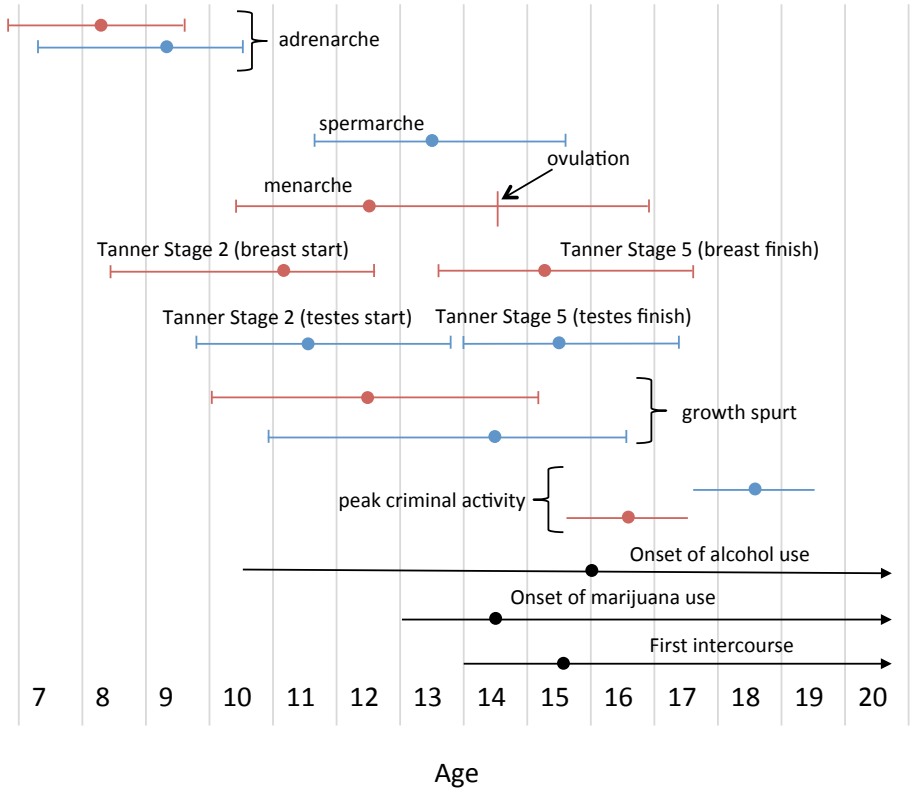
introduction will first provide a brief overview of the theoretical models and major empirical findings relevant to all three studies. Results from each study will then be presented. The concluding chapter will integrate findings from each study, highlighting common strengths, limitations, and future directions.

DEFINING PUBERTY

Puberty is generally defined as the hormonally-driven process of reaching reproductive maturity. The physical changes of puberty are driven by two biological processes, adrenarche and gonadarche. *Adrenarche* refers to the activation of the hypothalamic-pituitary-adrenal (HPA) axis and the secretion of androgens (DHEA, DHEA-sulphate, and androstenedione) from the adrenal glands, typically occurring around age 8 in girls and age 9 in boys (Blakemore et al., 2010; Dorn & Biro, 2011). In both boys and girls, levels of DHEA rise gradually and lead to body odor, pubic hair growth, and acne (Grumbach, 2002). Adrenal androgens continue to rise until the early 20s. *Gonadarche* refers to the activation of the hypothalamic-pituitary-gonadal (HPG) axis, the secretion of steroid hormones by the gonads (ovaries and testes), and ultimately gametogenesis (reproductive ability). The primary steroid hormone is testosterone in boys and estradiol in girls. These hormones ultimately lead to breast development and menarche in girls and voice changes and facial hair in boys. In girls, gonadarche typically begins between ages 8 and 14 (average age = 11; Blakemore, Burnett, & Dahl 2010) and is complete between 13 and 18 (average age = 15; Dyk, 1993; Marshall & Tanner, 1969; Strasburger & Brown, 1991). The average of age of menarche is 12 (McClintock & Herdt, 1996), and ovulation typically occurs 2-3 years after menarche (Apter, 1980). In boys, gonadarche typically begins between ages 9 and 15 (average age = 12) and is complete between ages 13 and 17 (Blakemore, Burnett, & Dahl 2010).

Pubertal development is often assessed using Tanner stages (Marshall & Tanner, 1969), which use external signs of puberty (e.g., breast development and pubic growth hair for girls, genital development and pubic hair growth for boys) as indices of maturation. Five stages are used, with the final stage representing reproductive maturity. Girls typically reach Tanner Stage 5 for breast development at age 15 (Marshall & Tanner, 1969; age range between 11-19). The upper half of Figure 1 (adapted from Hollenstein & Lougheed, 2013) shows a timeline of when these changed typically occur.

Figure 1. Timeline of Pubertal Changes and Trends for Initiation of Sexual Activity, Substance Use, and Delinquency.



These external physical changes are driven by gonadal and adrenal hormones, and levels of these circulating hormones increase with advancing pubertal development. In

boys, DHEA doubles over the course of adolescence. Boys experience dramatic increases in testosterone into young adulthood, with peak levels around age 18-19; these increases closely parallel pubertal stage (Rowe et al., 2004; Biro, Lucky, Huster, & Morrison, 1995; Granger, Schwartz, Booth, & Arentz, 1999; Shirtcliff et al., 2009). Testosterone levels also increase in girls, to a much smaller extent, and the association between testosterone and female physical development is less straightforward (Shirtcliff et al., 2009). Estradiol levels in girls increase markedly from pre- to post-puberty (Shirtcliff et al., 2009). Despite the clear functional relationship between hormonal increases and physical development, levels of pubertal hormones do not perfectly correlate with pubertal stage (Dorn & Biro, 2011): adolescents at the same Tanner stage may differ with regards to hormonal levels, and the hormonal ranges for each stage overlap. Previous studies have found moderate correlations between levels of pubertal hormones – particularly testosterone – and pubertal stage for boys, whereas correlations between pubertal stage, testosterone, and estradiol are minimal for girls (e.g., Shirtcliff et al., 2009; de Water, Braams, Crone, & Peper, 2013).

PUBERTAL STATUS VERSUS PUBERTAL TIMING

Two distinct but related aspects of pubertal development are relevant to risk-taking: pubertal status and pubertal timing. Pubertal status refers to the "absolute" level of physical development, with the highest level representing reproductive maturity. As summarized above, virtually all individuals experience hormonal changes, develop secondary sex characteristics, develop sexual attraction, and become reproductively mature during the second decade in life. These largely universal biological changes may contribute to overall (population-level) age trends in risk-taking. For example, post-pubertal adolescents are clearly more likely to engage in risky sexual activity – indeed, in

any sexual activity – then pre- or peri-pubertal adolescents. Changes in pubertal status also parallel age trends in other risk behaviors. Among boys, the pronounced increase in testosterone throughout adolescence and the peak in early adulthood parallel the developmental course of risk behavior (Rowe et al., 2004, Mazur & Booth, 1998). Gender differences in the biological changes at puberty may also explain gender differences in risk-taking that emerge during this period. For example, males undergo puberty later in life, and delinquency peaks earlier for girls than for boys (US Department of Justice, 2003). Thus, changes in pubertal status may explain why, on average, adolescents show higher levels of risk behavior (and initiate risk behavior at this time), and why this trend is particularly pronounced in males. Because changes in pubertal status are largely universal and because many of these changes are complete by mid-adolescence (e.g., menarche, pubic hair development), studies on pubertal status and its associations with risk-taking (and other psychological outcomes) are answering questions that are limited to a specific age range—the age when pubertal status is changing.

In contrast, *pubertal timing* refers to how developed an individual is relative to his or her same age peers. There is considerable variability in the timing of pubertal changes, and much of this variability is due to genetic influence (Rowe, 2000; Ge, Natsuaki, Neiderhiser, & Reiss, 2007). Pubertal timing is an individual difference variable with both transient and persistent effects, as the timing of puberty may set in motion a cascade of experiences culminating in important developmental outcomes. The *life course perspective* (Elder, 1998) is useful framework for in understanding this developmental cascade. This perspective emphasizes how the timing and context of discrete life transitions have long-term implications. For example, Cavanagh, Riegle-Crumb, and Crosnoe (2007) applied a life course framework to study the connection between early age at menarche and poor educational outcomes. They found that girls with early

menarche were more likely to fail courses at the start of high school, and subsequently more likely to drop out. Those who stayed in high school had lower grades upon graduation. These educational outcomes during adolescence, in turn, have implications for later education (e.g., college matriculation) and for socioeconomic attainment throughout the lifespan. The proposition that early timing can set adolescents on a negative trajectory also applies to other outcomes in young adulthood, including depression, teenage motherhood, and risk for intimate partner violence (through a process of "subjective weathering," reviewed in Foster, Hagan, & Brooks-Gunn, 2008). Thus, pubertal timing, moreso than pubertal status, is relevant to both the time-limited experience of puberty *and* to risk for future negative outcomes in late adolescence and adulthood.

Pubertal status and pubertal timing can be hard to distinguish empirically. To identify the unique effects of pubertal status, investigators commonly control for chronological age; however, pubertal status relative to age is essentially a measure of pubertal timing. However, in examining associations between puberty and risk-taking, these are theoretically distinct: Pubertal status is a potential mediator of age trends in risky behavior—why adolescents, on average, engage in riskier behavior than children or adults. Pubertal timing, in contrast, is an individual differences variable that may explain why some adolescents are more at risk than others.

PUBERTY AND RISK-TAKING

Decades of research on pubertal timing have yielded several robust findings. First, adolescent girls with early pubertal timing are more likely to engage in risky behavior, including substance use and delinquency, than later maturing girls (Caspi & Moffitt, 1991; Ge et al. 1996, Mendle et al., 2007, Graber et al., 1997). On average, early

maturing girls initiate substance use earlier, drink more frequently, and show higher levels of substance use disorders in young adulthood (Wiesner et al., 2002; Wilson et al., 1994, Dick, Rose, Viken, & Kaprio, 2000). Less is known about the impact of pubertal timing in boys, but there is growing evidence for a similar trend, with early maturing boys more likely to engage in a range of risk-taking behaviors at an earlier age (Mendle & Ferrero, 2012; Biehl, Natsuaki, & Ge, 2007; Costello, Sung, Worthman, & Angold, 2007; Udry, 1991; Wichstrom, 2001). Some research has found higher rates of alcohol use among both early and late maturing boys (Anderson & Magnusson, 1990). Second, the association between pubertal timing and risky behavior is, at least in part, mediated and moderated by the social context in which these biological changes unfold (Stattin & Magnusson, 1990; Ge et al., 2002; Haynie, 2003; Cavanagh, 2004). Specifically, peer relationships and peer group characteristics have been identified as key mediators of this association (Haynie, 2003; Stattin & Magnusson, 1990).

There are many theories for the associations between puberty and adolescent risk-taking. A common theme underlying all of these theories is a *maturity gap*. For example, there are gaps between the age at which adolescents are capable of reproduction and when society considers them to be "adults," a gap between maturation of the brain's socioemotional and cognitive control systems, a gap between exposure and access to drugs, alcohol, sexual partners and the psychological capacity to abstain from and/or cope with these potential stressors. All of these gaps have been proposed as explanations for risk-taking in adolescence, and all are wider for adolescents with early pubertal timing. Below, I describe these various "maturity gap" theories in greater detail, divided according to whether they primarily emphasize biological versus social processes.

Biological Theories

Raging hormones There is a widely-held lay assumption that teenagers take risks because of "raging hormones," (Buchanan, Eccles, & Becker, 1992), and this view has at least some support from animal and human research. In addition to driving the development of secondary sex characteristics and reproductive maturity, gonadal hormones have both organizational and activational effects on the adolescent brain (Forbes & Dahl, 2010; Kuhn et al., 2010; Blakemore, Burnett, & Dahl, 2010), which manifest in behavioral changes. In the Syrian hamster model, for example, testosterone released at puberty increases sexual and aggressive behavior (Schulz & Sisk, 2006) and reduces aversion to novel stimuli (Primus & Kellogg, 1989) in males. In humans, individual differences in concentrations for pubertal hormones (testosterone and estradiol) are associated with both internalizing and externalizing in males and females (Reynolds, Tarter, Kirisci, Kirillova, Brown et al., 2007; Eriksson, Kaprio, Pulkkinen, & Rose, 2005; Costello et al., 2007; de Water et al., 2013). Moreover, sensation-seeking—the tendency to seek out high-intensity, novel, exciting experiences (Zuckerman, 1994)—is associated with pubertal development, controlling for chronological age (Kretsch & Harden, 2013; Forbes et al., 2010; Martin et al., 2002; Steinberg et al., 2008; Zuckerman, Buschbaum, & Murphy, 1980).

Gonadal hormones have been linked to remodeling of reward-related regions including the striatum, nucleus accumbens, and dopaminergic pathways to the prefrontal cortex (Sisk & Zehr, 2005). These structural effects are thought to underlie age changes in reward sensitivity, sensation-seeking, and reward-motivated behavior. The development of the neural structures underlying these "hot" affective processes may be more closely tied to pubertal development than to chronological age (Steinberg, 2008; Nelson et al., 2005; Martin, et al., 2002; Zuckerman et al., 1980). The hormonal "surge"

at puberty occurs before maturation of the brain regions governing self-regulation and impulse control (Steinberg, 2008; Nelson et al., 2005; Casey, Getz, & Galvan, 2008; Dahl, 2004), rendering adolescents—particularly early maturing adolescents—more sensitive to the rewarding aspects of risky behavior without the capacity for inhibitory cognitive control. The perspective that gonadal and adrenal hormones play a critical role in the link between puberty and risk-taking aligns with research implicating these hormones in a range of risk-related phenotypes, not only sensation-seeking but also social dominance (Tarter et al., 2007), delinquency (Rowe, Maughan, Worthman, Costello, & Angold, 2004), substance use (Reynolds et al., 2007), and early sex (Halpern, Udry, & Suchindran, 1998).

However, the association between pubertal hormones and risk-taking is not simple and not direct. Animal studies have produced inconsistent results. Vetter-O'Hagen & Spear (2012) examined novelty-seeking behavior (time spent sniffing at novel stimuli) in adolescent and adult rats. Puberty was measured by post-mortal hormone (testosterone, estradiol, and progesterone) levels and by genital inspection. Controlling for age, pubertal status was not associated with novelty seeking behavior. In humans, studies linking testosterone with delinquency have been inconsistent (reviewed in Duke, Balzer, & Steinbeck, 2014). Hormonal influences are moderated by social context (Rowe et al., 2004) and by other hormones (Mehta & Josephs, 2010). Thus, while endocrine changes are the driving force of pubertal development, they do not fully explain the association between puberty and adolescent risk-taking.

Genetic factors. Additional evidence for the role of biological processes in the association between puberty and risk-taking comes from behavioral genetic research. Research using twins has shown that pubertal timing and levels of pubertal hormones are both heritable (Ge et al., 2007; Mustanski, Viken, Kaprio, Pulkkinen, & Rose, 2004).

Behavior genetic studies of age at menarche have yielded heritability estimates between 60% and 70%. Genetics also influence timing of breast development and skin changes (Ge, et al., 2007; Mustanski et al., 2004). Level of testosterone is 40-70% heritable in adolescence, with some studies suggesting differences in heritability by gender (Hoekstra, Bartels, and Boomsma, 2006; Harden, Kretsch, Tackett, & Tucker-Drob, 2014; Koenis et al., 2013). Likewise, genetic influences on numerous forms of risk-taking are apparent in adolescence (Koopmans & Boomsa, 1997; Kendler, Schmidt, Aggen, & Prescott, 2008). There is some evidence that the association between puberty and risk-taking is at least partly due to common genetic effects, i.e., that the same genetic factors that predispose one for early maturation also promote risky behavior (Harden & Mendle, 2012). For example, the short allele of the X-linked androgen receptor gene (AR1) has been associated with both aggression and impulsivity in males and with earlier ages of physical maturation in females (Comings et al., 2002).

Psychosocial Theories

Although genetic and hormonal differences are likely important in explaining the relation between puberty and risk-taking, most developmental research has emphasized psychosocial processes. One widely promoted theory is the *maturity disparity hypothesis* (reviewed in Ge & Natsuaki, 2009), which focuses on the gap between physical and psychological development. A more mature physical appearance instigates a number of social changes, including expectations of romantic involvement and more freedom from parents. Adolescents who face these transitions at a younger age are less psychologically mature and may lack the cognitive and emotional resources necessary to cope with the stress associated with these changes (Caspi & Moffit, 1993). A related theory is that individuals become physically mature (capable of reproduction) well before society

considers them to be adults. Risky behaviors including substance use and delinquency may reflect adolescents' response to this gap between physical and social maturity (Moffitt, 1993).

Another line of research has focused on the interaction between the social environment and physical development. The peer group is a key player in the association between pubertal timing and risk-taking. Adolescents who mature earlier are more likely to affiliate with older peers, who provide access and opportunity to engage in risky behavior (Haynie, 2003; Stattin and Magnusson, 1990; Halpern et al., 2007). Adolescents who appear more physically mature may also be granted more autonomy from parents than their same-age counterparts, which may allow more opportunities for risk-taking. The *contextual amplification hypothesis* (Ge et al., 2002; Stattin et al., 2011) posits an interaction between pubertal timing and contextual influences on risk-taking such that early maturation leads to increased risk-taking only in the presence of other environmental risk factors. Consistent with this hypothesis, one study found that early maturing girls are more prone to delinquency in co-educational schools, but not in all-girls schools (Caspi, Lynam, Moffitt, & Silva, 1993). Other research has shown that the effects of early timing are more pronounced in the context of low parental monitoring (Mrug et al., 2008) and neighborhood disadvantage (Obeidallah, Brennan, Brooks-Gunn, & Earls, 2004).

MEASURING PUBERTAL DEVELOPMENT

How researchers measure these processes is a critical factor in integrating findings from the vast corpus of literature on psychosocial sequelae of pubertal development. Two factors in particular have determined how puberty is measured: What is feasible given the research setting and what is appropriate given the research question.

In obtaining an objective measure of pubertal status, Tanner staging by a trained clinician (typically a physician or nurse) is considered a "gold standard" (Dorn, Dahl, Woodward, & Biro, 2006). These measures do not always align with an adolescent's perception of these changes. Because it is often impractical to include physical examinations in non-clinical research settings, much of the research on puberty has relied on adolescent self-report. The most widely used self-report tool is the Pubertal Development Scale (PDS; Peterson, Crockett, Richards, & Boxer, 1988), which asks adolescents about whether specific physical changes have started, are underway, or have completed. Self-reports on the PDS, however, are only moderately correlated with Tanner stage ratings from physical examination (Shirtcliff, Dahl, & Pollak, 2009; Dorn et al., 2006; Brooks-Gunn et al., 1987). Self-reported age at menarche is a reliable indicator of pubertal timing in girls (Bean, Leeper, Wallace, Sherman, & Jagger (1979), but this captures only one aspect of pubertal development. Moreover, there is no equivalent, widely used measure in boys. Thus, the potential unreliability of self-report measures of puberty has been of concern, and these measures have been criticized as poor indicators of actual physical changes (Dorn et al., 2006). However, while self-report measures may not perfectly measure actual physical changes, they still predict important health outcomes. In fact, some of the largest effect sizes have been found in studies using self-report measures (e.g., Deppen et al., 2012; Lanza & Collins, 2002; Rierdan et al., 1988; Michael & Eccles, 2003). This suggests that perceived development is a meaningful construct itself, one which is likely influenced by both biological and contextual factors. The studies in this project presuppose that different measures of puberty—hormonal, self-reported physical development, age at menarche, and peer comparisons—are relevant to different research questions.

PUBERTY IN CONTEXT

Puberty is the quintessential example of a biopsychosocial phenomenon. Families, friends, schools, neighborhoods, and cultures are all important in understanding the impact of this transition. The studies in this project all make use of nested or clustered data—individuals clustered within schools or within families. This method violates the assumption of independence of observations, and statistical adjustments must be made in analyses to deal with dependencies. In this way, nested data is sometimes considered a "nuisance," a trade-off between efficiency of sampling and efficiency of analysis. An alternate perspective is that this non-independence is theoretically meaningful and that nested data allows us to measure multilevel effects. Two forms of nested data are used in this project: (1) peer network data, which involves individuals nested in schools and in peer groups within those schools, and (2) behavior genetic data, which involve siblings nested in families.

Social network data is increasingly common within developmental research, as researchers have developed sophisticated methods for modeling how social networks form and change within a school. Peers play a key role in nearly all theories of pubertal timing and risk-taking. It is not surprising, therefore, that a number of studies on pubertal timing have drawn on network data (Haynie, 2003; Cavanagh 2004; Jaccard et al., 2005). A major benefit of these datasets is that they provide peer-reports of peer behavior, rather than relying on adolescents to report on the behavior of their friends. Direct report of friends' pubertal timing offers a valuable (and rare) chance to explore how social comparisons and peer group composition shape adolescents' perceptions of their own pubertal development.

Behavior genetic designs, which have historically been used to estimate heritability, are also useful in studying environmental influences. A sibling-comparison

or twin-comparison design offers a quasi-experimental approach to testing hypotheses about environmental influence. This approach tests whether twins who are discordant for exposure to an environmental variable are also discordant for a behavioral outcome (Lahey & D'Onofrio, 2010). For example, the second study in this project applies a discordant twin design to social network data, in order to test whether twins who differ in the delinquency of their peer groups also differ in their own delinquency, and whether this within-twin effect is moderated by pubertal development.

Twin studies are also useful in identifying endophenotypes—biologically-based intermediary phenotypes that bridge genetic risk with higher order, more complex phenotypes (Gottesman & Gould, 2003). An endophenotype is a characteristic that is heritable and mediates genetic influence on a phenotype, and therefore can be used as an index of an individual's genetic risk. Identifying endophenotypes for psychological disorders is important as the pathways between specific genetic variants and behavioral outcomes remain elusive in the current "post-genomic era" (following the completion of the Human Genome Project). Quantitative genetic studies (i.e., twin and sibling studies) do not identify specific genetic variants linked with behavioral outcomes, but they can shed some light on the pathway from gene to phenotype.

OVERVIEW OF STUDIES

This project includes three studies. Data were drawn from two sources: the Texas Twin Project (TXT; Study 1) and the National Longitudinal Study of Adolescent Health (Add Health; Studies 2 and 3). Major strengths of the TXT data include its ethnic diversity, the fact that it is genetically informative, and that it includes hormonal measures, behavioral, and self-report measures. Major strengths of the Add Health data are its size, nationally representative sample, genetically informative data, and social

network data. These studies integrate several lines of research from sociology, behavioral genetics, and developmental, social, clinical, and cognitive psychology. Study 1 examined the cognitive processes underlying propensity for risk-taking, tested whether these processes were linked to pubertal hormones and whether hormones mediated genetic influence on substance use. Study 2 examined how perceived physical development moderated peer influence on risk-taking, using a quasi-experimental approach to disentangle selection and socialization. Finally, Study 3 examined how different measures of perceived pubertal timing might be shaped by peer comparisons.

Chapter 2: Testosterone and Risk-Taking

BACKGROUND

Adolescence is a time of heightened involvement in a number of risky behaviors, including initiation of alcohol and drug use (Johnston, O'Malley, Bachman, & Schulenberg, 2013; CDC, 2009). Substance use in adolescence is associated with both immediate and long-term negative health and academic outcomes. Individuals who start drinking at an earlier age are more likely to develop alcohol-related problems later in life (Grant et al., 2005). Although substance use is widespread in adolescence, there is substantial variation in the extent of substance use in this age group, which is in part due to genetic influence. Genetic influences on substance use and other forms of risk-taking become apparent during adolescence (Koopmans & Boomsa, 1997; Kendler & Schmidt, 2010), yet the specific genetic underpinnings of substance use and risk-taking in adolescence remain elusive.

Pathways from Genes to Behavior

Understanding how genes influence complex behavioral phenotypes is a challenge, and one promising approach is to identify endophenotypes—intermediate, measurable phenotypes that link genetic risk to more complex behavioral phenotypes such as substance use (Gottesman & Gould, 2003). An endophenotype is a characteristic that a) is heritable, b) is associated with a phenotype (such as substance use), and c) mediates genetic influence on a phenotype. One candidate endophenotype for adolescent risk-taking is testosterone, a gonadal sex steroid. Levels of testosterone increase

dramatically during adolescence, doubling in females and increasing nearly tenfold in males (Biro, Lucky, Huster, & Morrison, 1995; Granger, Schwartz, Booth, & Arentz, 1999). In males, release of testosterone by the testes stimulates voice changes and the growth of body hair. In females, testosterone is released by the adrenal gland and triggers the growth of pubic hair (Grumbach & Styne, 2003). In addition to causing morphological changes, testosterone also affects the adolescent brain, by acting on existing neural networks (activational effects) and by causing structural changes (organizational effects) (Witt, 2008; Schulz, Molenda-Figueria, &Sisk, 2009). Until recently it was assumed that organizational effects of testosteorone were limited to the prenatal period; however, there is growing evidence that the brain undergoes structural changes at puberty, and testosterone organizational effects are also apparent at puberty. In particular, testosterone has structural and functional effects on the dopaminergic neural networks involved in reward processing, including the ventral striatum and its connections to the prefrontal cortex (Witt, 2008; Peper & Dahl, 2013; Forbes & Dahl, 2010).

Although testosterone and other pubertal hormones increase in nearly all individuals, there are individual differences in the timing, the magnitude, and the stability of endocrine changes at puberty, and genes partially account for this variation (reviewed by Caramaschi, Booij, Petitclerc, Boivin, & Tremblay, 2012). Moreover, individual differences in endogenous testosterone are associated with a range of risk behaviors, including substance use (Reynolds et al., 2007; Eriksson et al., 2005; Costello et al.,

2007; de Water et al., 2013). Testosterone therefore appears to satisfy two conditions that make it a viable candidate endophenotype for risk-taking.

Testosterone and the Reward System

Testosterone might influence risk-taking through its impact on the reward system. According to the dual systems model of adolescent neurodevelopment, adolescents' heightened propensity for risk-taking (relative to adults and children) results from the asynchronous development of two neural systems (Steinberg, 2008; Casey, Getz, & Galvan, 2008). The socioemotional system, which governs affective or "hot" cognitive processes such as reward-sensitivity, behavioral activation, and sensation-seeking, matures relatively rapidly in early adolescence, before the development of brain regions involved in self-regulation and cognitive control ("cold" processes), which do not mature fully until adulthood. When faced with an opportunity to engage in risky behavior, adolescents are sensitive to the potential rewards of risky behavior, while their capacity for impulse control is relatively immature.

There is some evidence that the socioemotional, "hot" affective system is more closely tied to pubertal development than to chronological age (Steinberg, 2008; Nelson et al., 2005; Martin et al., 2002; Zuckerman et al., 1980), and testosterone, in particular, may act on the socioemotional system to enhance reward sensitivity and increase reward-seeking behavior. Animal and human studies have provided convergent support for testosterone's role in the reward system in general and, more specifically, in the reinforcing properties of alcohol and drugs (Wood, 2004; Witt, 2008). For example, rats that are administered testosterone show increased lever-pressing to obtain drugs (Clark,

Lindenfeld, & Gibbons, 1996) and increased conditioned place-preference (Alexander, Packard, & Hines, 1994). Apter and Eriksson (2003) found that rats that were selectively bred for alcohol preference showed higher basal serum testosterone and a higher rates of testosterone elevations in response to alcohol compared to a control group. In humans, testosterone is associated with sensation-seeking, a trait that is related to reward sensitivity and risk-taking (Campbell et al., 2010). Laboratory studies with humans have demonstrated structural and functional associations between testosterone and reward processing. Male and female adolescents with higher salivary testosterone showed enhanced activation of the ventral striatum in anticipation of and response to reward (Forbes et al., 2010; Op de Macks et al., 2011). One recent study found higher levels of testosterone in young adults were correlated with reduced integrity of the frontostriatal white matter tracts, and that this reduced integrity predicted preference for immediate (vs. delayed) rewards (Peper et al., 2013).

Testosterone and Risk-Taking

Testosterone's effects on reward and incentive processing are thought to translate into risk-taking behavior. Higher testosterone levels predict laboratory measures of risk-taking in both males and females, including performance on the Balloon Analogue Risk Task (BART; Peper, Koolshijn, & Crone, 2013) and on the Iowa Gambling Task (IGT; Stanton, Liening, and Schultheiss, 2011). In their study of testosterone and risk-taking, Peper, Koolshijn, and Crone (2013) found that grey matter volume in the medial orbitofrontal cortex mediated associations between testosterone and risk-taking on the BART. Similarly, Stanton, Liening, and Schultheiss (2011) found that higher salivary

testosterone was associated with increased reward seeking and decreased harm avoidance on the Iowa Gambling Task (IGT). These findings paralleled an early study in which young adult women who were administered testosterone showed a similar pattern of higher reward seeking and lower harm avoidance on the IGT (Van Honk et al., 2004).

Laboratory-based paradigms such as the IGT (Bechara, 2007) and the BART = (Lejuez, 2002) are useful because they reveal quantifiable differences in potential neurobehavioral endophenotypes, and they overcome limitations of self-report measures. Using the IGT, for example, one can measure changes in approach and avoidance behavior over time and distinguish between harm avoidance and reward sensitivity (e.g., Cauffman et al., 2012). Most laboratory research on the cognitive processes underlying adolescent risk behavior has relied on single tasks to distinguish constructs of interest. This is a limitation, however, because behavioral measures show a great deal of task-specific variance, and there is often little overlap between self-report measures and behavioral measures that are designed to measure similar constructs (Braams et al., in press; Reynolds et al., 2006; Cyders & Coskunpinar, 2011). Few studies have specifically examined how laboratory measures of decision-making overlap with each other, with self-report measures, and with general cognitive ability.

Goals of the Current Study

The overall goal of the current study was to examine testosterone and reward seeking as potential endophenotypes for substance use. We examined associations between salivary testosterone, reward seeking, and initiation of alcohol and marijuana use in a sample of adolescent twins. We measured reward seeking using a battery of self-

report and behavioral tasks. Guided by the dual systems model (Steinberg et al., 2008; Harden & Tucker-Drob, 2011; Casey, Getz, & Galvan, 2008), we expected that this multidimensional approach would crystallize on at least two components: one reflecting the "cold" inhibitory processes of the cognitive control system, and one reflecting "hot" affective processes of the socioemotional system (reward-seeking and/or sensation-seeking). We hypothesized that testosterone would predict substance use initiation via reward seeking, controlling for chronological age, and that associations would be greater for males than for females. We used a twin design to test the extent to which observed associations between testosterone, reward seeking, and substance use initiation were due to common genetic factors. We predicted that testosterone would mediate genetic influence on substance use, more so in males than in females.

METHOD

Participants

Adolescent twin pairs were recruited from an ongoing study of twins in a large metropolitan area in the southwestern United States. Twins were identified from school rosters. The current analyses include data from 530 individuals total, ages 13-20 (M_{age} = 16.0, *SD* = 1.5). Race and ethnicity were assessed by self-report. Sixty-four percent of the sample was non-Hispanic White, 19% was Hispanic/Latino, 12% was African American, and 5% was of another race/ethnicity. Socioeconomic status was operationalized as the highest level of education attained by the parent(s), ranging from 1=*1st Grade* to 22=*Professional degree after bachelors degree*. This was assessed by parent report (*M* = 18.27, *SD* = 2.80, median=18=*Bachelor's degree*).

Measures

Zygosity. All opposite-sex pairs were classified as dizygotic (DZ). Zygosity for same-sex twin pairs was determined from parent and adolescent responses to a questionnaire about perceived similarity. The questionnaire included items assessing how often twins were mistaken for one another and how similar they were in facial features, hair, and eye color. Previous studies have validated the accuracy of these measures with molecular genetic zygosity classification (Rietveld, van der Valk, Bongers, Stroet, Slagboom, & Boomsma, 2000). For each pair of twins, the parent and twin scores on zygosity items were subjected to a latent class analysis (LCA). The fit of two-class LCA model was assessed using entropy, with values closer to 1.0 indicating better classification. The LCA model had good entropy (.992), indicating very little uncertainty in classification. Final zygosity classifications were as follows: 96 monozygotic (MZ) pairs (52 female; 44 male) and 165 DZ pairs (39 female, 50 male, and 76 opposite-sex). For families with triplets, two members of the triplet set were selected (and data from the third triplet omitted); selection prioritized same-sex pairs and participants with complete data.

Testosterone. Saliva samples of approximately 2-3 mL were collected by passive drool, following recommendations by Granger, Shirtcliff, Booth, Kivlighan, and Schwartz (2004). Samples were taken from both twins at the same time at the beginning of the testing sessions. Samples were taken at one of three appointment times: 9:00 a.m. (27%), 12:00 pm (47%), and 3:00 pm (26%). Participants were instructed to avoid eating or drinking for 2 hours prior to their appointment time. They were also instructed not to

floss on the morning of the appointment and not to smoke for four hours before the appointment time. Whenever possible, females were within the first 14 days of their menstrual cycle at the assessment. All samples were frozen at -40° C until they were shipped on dry ice to Dr. Clemens Kirschbaum's laboratory at the Technical University of Dresden for analyses of hormone levels. After thawing, samples were centrifuged and assayed using commercially available chemiluminescence-immunoassays with high sensitivity. Eight participants were missing testosterone because they declined to provide a sample, the sample was contaminated, the vial was defective, and/or they did not provide enough saliva to be assayed.

Testosterone concentrations were inspected for outliers. One female had an abnormally high value (>700 pg/mL), which was winsorized to 3 standard deviations above the mean concentration in females. Testosterone concentrations were log-transformed to approximate a normal distribution. Data were corrected for time at sample collection, hours since waking, oral contraceptive use (females only), menarcheal status (pre vs. post, females only), steroid use, and menstrual cycle phase (whether female participant was within 2 weeks of menstrual cycle). Data were then residualized by age and age-squared and standardized by sex into z-scores.

Pubertal Timing. Pubertal timing was assessed using the Pubertal Development Scale (PDS, Petersen Petersen, Crockett, Richards, & Boxer, 1988), a widely used self-report measure of pubertal development. The PDS includes items asking about changes in skin, height, underarm and pubic hair, breast development (for females), voice changes (for males), and facial hair (for males). Participants rated these items on a 4-point scale

ranging from 1=*Not Yet Begun to Change* to 4=*Finished Changing*. Females were asked if they have begun menstruating (95% of females in the sample reported that they had). The question about menstruation was recoded to be consistent with the response scale of the other items (*No*=1, *Yes*=4). Pubertal status was calculated by averaging these items (range = 1 to 4; $M = 3.17$, $SD = 0.63$). Females reported more advanced development ($M = 3.45$, $SD = .49$) than males ($M = 2.90$, $SD = .63$). Mean scores were standardized by age and sex to obtain a measure of pubertal timing. Pubertal timing was positively correlated with testosterone in males ($r = .22$, $p < .01$) but not in females ($r = -.01$, $p = .83$).

Substance use. Substance use was assessed by self-report. Three dichotomous measures were used. Adolescents were asked if they had ever had a drink containing alcohol outside the presence of their parents. Twenty-four percent (24%) of adolescents reported ever using alcohol outside their parents' presence; these were classified as having initiated alcohol use (Alcohol Use = 1). Those who endorsed this item were also asked how often in the past year they had "gotten drunk or very high" on alcohol. Seventeen percent of all adolescents endorsed ever being drunk in the past year (Past year intoxication = 1). Adolescents were also asked whether they had ever used marijuana (18% reported yes), and this was also analyzed as a dichotomous outcome variable.

Self Reported Impulsivity. Impulsivity was measured using the UPPS impulsivity scale (Whiteside & Lynam, 2001). This scale includes 45 items, rated on a 4-point scale (1 = "Disagree strongly," 4 = Agree strongly"). It distinguishes between four facets of impulsivity. *Sensation-seeking,* the tendency to pursue novel and exciting experiences (Zuckerman, 1994), was assessed with 12 items ($\alpha = .84$) such as "I

sometimes like doing things that are a bit frightening" (*M* = 2.87, *SD* = .55). *Premeditation*, the tendency to think about potential consequences before acting, was assessed with 11 items (α = .83) such as "I usually think carefully before doing anything" (*M* = 2.92, *SD* = .48). *Perseverance,* the ability to remain focused on a task that may be difficult, was assessed with 10 items (α = .55) including "Once I start a project, I almost always finish it" (*M* = 2.99, *SD* = .52). *Urgency*, the tendency to experience strong impulses under conditions of negative affect (Whiteside & Lynam, 2001) was assessed with 12 items (α = .80) such as "It is hard for me to resist acting on my feelings" (*M* = 2.19, *SD* = .56).

Planning. Planning was assessed using the Tower of London (TOL; Shallice, 1982), a test of strategic planning. This task included an apparatus (which appears on the computer screen) that had three rods, each of which held a number of balls. The tallest rod could hold three balls, the next tallest rod could hold two balls, and the shortest could hold one ball. On each trial, a participant was shown an image of the "goal instrument," which included a specific arrangement of balls on the rods. There was also a "game instrument," which included a different arrangement of balls. The goal was to rearrange the balls on the game instrument to match the goal instrument, using the fewest number of moves possible. The player used the mouse to drag and drop the balls from one rod to another. The task included five sets of four problems, starting with easy problems that could be solved in a minimum of three moves and progressing to difficult ones that required a minimum of seven moves. Errors and excess moves on this task can reflect poor planning ability, poor working memory, and/or poor inhibitory control (Albert &

Steinberg, 2011). The primary outcome of interest in the current study was the total number of excess moves that a participant made across all the trials (range = 0-11, *M* = 2.60, *SD* = 2.31).

Risky Decision Making. Risky decision-making was measured using two laboratory tasks, both administered on desktop computers. The BART (Lejuez et al., 2002) is a measure of propensity for risk-taking. A balloon, a balloon pump, and a meter indicating cash earnings were displayed on the screen. The object of the game was to earn money by inflating a series of balloons as much as possible without having them pop. The player inflated the balloon by clicking the mouse, and for each pump, $.05 was added to the player's "temporary account." If the balloon exploded, the money in the temporary account was lost. At any point during the game, the player could choose to transfer money to the permanent account. The BART included 30 balloons presented one at a time. Balloons varied in the point at which they exploded. Players were told that explosions could occur as early as the first pump and as late as the time at which the balloon filled the computer screen. Players were told that their performance on this task, measured in total earnings, would go toward the bonus compensation they would receive at the end of the study. The main variable of interest was the average number of pumps per balloon, limited to balloons that did not pop ("adjusted average;" Lejuez et al., 2002). Higher scores represented greater risk-taking (*M* = 31.1, *SD* = 12.8).

The Stoplight Game (Steinberg et al., 2008) is a simulated driving task in which the player "drives" a car along a straight track, trying to reach a specified location in under 5 minutes. In this game, the car passed through 20 intersections, each of which had

a stoplight that cycled from green to yellow to red as the vehicle approached it. When the light turned yellow, the player decided whether to stop, which resulted in a short delay, or to go through the intersection and risk a "crash" with another vehicle, which resulted in a longer delay. The timing of the traffic signals and the probability of a crash were varied--at some intersections, not braking in time inevitably resulted in a crash, while at others, it was possible to drive through safely. Each intersection had three possible outcomes: 1) the player applied the brakes and safely stopped, 2) the player went through the intersection successfully, or 3) the player crashed into another car, which could result from either failure to brake at all or failure to brake in time to avoid the crash. The computer recorded the outcome at each intersection. Risky decision-making was indexed by the percentage of intersections where the player drove through the yellow light (M = .45, SD = 12.8).

Affective Decision-Making. Affective decision-making includes two types of implicit learning—*harm avoidance* and *reward sensitivity*. These two types of implicit learning were assessed using a modified version of the IGT (Bechara et al., 1994). The IGT measures affective decision-making under conditions of uncertainty. In this version of the game, the player tried to earn hypothetical money by "playing" or "passing" on cards. Four decks were presented on the screen, labeled A, B, C, and D. The screen displayed a running total of the player's money. In each trial, an arrow appeared pointing toward one of the four decks. The player had 4 seconds to decide whether to play or pass on a card from that deck. If they decided to play, the screen would show how much money they had won or lost, and the running total of earnings would be updated. If they

decided to pass, they received no feedback and the running total stayed the same. The task was administered in 6 blocks of 20 trials (120 trials total). Two of the decks were "bad," in that repeatedly drawing from these decks would result in a net loss of earnings. Two of the decks were "good," in that drawing cards from these decks would result in a net gain of earnings.

Two outcomes were of relevance to the current study's research questions: a) the extent to which a player learned to play from good decks, and b) the extent to which a player learned to avoid the bad decks. Using a mixed effects model that included test block (1-6) as a within subjects variable, we estimated a slope and intercept for good decks and bad decks to characterize each participant's performance over the six blocks. The slope indicated change in percentages of good and bad plays over the six blocks. The intercept was centered to indicate the percentages of good and bad plays on the final block. Intercepts and slopes were highly correlated (r = .96, p < .01). For the current analyses, we used the intercept values as indicators of reward sensitivity and harm avoidance. Higher intercept values for good and bad decks indicated higher percentages of good and bad plays made by the end of the task, respectively.

Delay Discounting. A Delay Discounting task was used to measure relative preference for immediate versus delayed reward. In this task, administered on computers, participants were asked to choose between two hypothetical monetary rewards—a smaller immediate reward (e.g., five dollars in 1 week) and a larger delayed reward (e.g., 1,000 dollars in one year). The task included six blocks with nine choices presented in each block. The delayed outcome was held constant at $1,000, the starting value of the

immediate reward was $200, $500, or $800, and the time to delay was 1 day, 1 week, 1 month, 3 months, 6 months, or 1 year, all of which were presented in random order. Each block started with a choice between a delayed reward of $1,000 and a smaller (i.e., $200, $500, or $800) immediate reward. If the participant chose the delayed reward, the next choice raised the value of the immediate reward. If the participant chose the immediate reward, the next choice lowered the value of the immediate reward. A total of nine choices were presented in each block, and the participant's responses converged at a point when the preference of immediate reward and delayed reward was equal. The outcome of interest was the discounting rate—the rate at which participants discounted the value of a fixed reward (e.g., $1,000) depending on the length of delay. Following previous research (e.g., Steinberg et al., 2009), we used the hyperbolic delay function to estimate the steepness of the discounting rate. A steeper (higher) discount rate reflected a greater preference for immediate rewards.

Intelligence. Intelligence was assessed using the Wechsler Abbreviated Scale of Intelligence (WASI; Psychological Corporation, 1999). The WASI can be administered in approximately 30 minutes and includes four subtests (Vocabulary, Matrix Reasoning, Block Design, and Similarities). The WASI is highly correlated with the Wechsler Adult Intelligence Scale and the Wechsler Intelligence Scale for Children (Wechsler, 1999). It has been age-normed for individuals between ages 6 and 98, yielding a standardized measure of Full-Scale IQ (FSIQ). The mean IQ in the sample was 103.2 (*SD* = 13.2), which is consistent with population norms.

Procedure

Twin families were identified from school records and were sent a recruitment letter inviting them to participate in the study. The recruitment letter offered compensation of "up to $50" per twin ($100 total for each family). The letter instructed interested participants to sign up online or by calling the lab. Research assistants called families who signed up online, explained the study, answered questions, and scheduled an appointment. Female participants were asked the date of their last menstrual cycle, and, whenever possible, appointments were scheduled within two weeks of the start of their menstrual cycle. Families were instructed to bring signed parent consent forms to the appointment.

All data collection occurred at the laboratory at the University of Texas at Austin. Procedures were approved by the Institutional Review Board at the University of Texas at Austin. When participants arrived, they handed in signed parent consent forms and completed assent forms. Research assistants took saliva samples. Participants filled out the survey on computers in private rooms. Following the survey, the WASI was administered by the research assistant. Finally, the computer tasks were administered. The entire protocol took an average of three hours. Each twin was assessed in a separate room by a different research assistant. To motivate participants on the tasks, they were told that their task performance would influence their compensation. However, regardless of their performance on the tasks, all adolescents were compensated $50 for their participation.

Analytic Plan

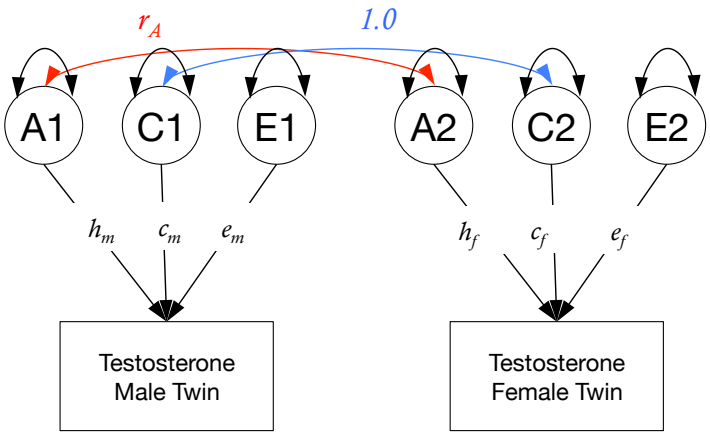
Analyses were performed using a structural equation modeling (SEM) framework in M*plus* version 7.0, using Full Information Maximum Likelihood (FIML) to account for missing data. Absolute model fit was assessed using the using the root mean square error of approximation (RMSEA; Steiger, 1990). RMSEA values of less than 0.05 indicate a close fit to the data, and values up to 0.08 represent reasonable errors of approximation. Differences in model fit were assessed using differences in log-likelihood and in chi-square, which both show a chi-square distribution.

In the current study, we were primarily interested in associations among three phenotypes: testosterone, reward seeking, and substance use. Analyses were performed in three main steps. First, we assessed genetic and environmental variance in testosterone using a twin model. The twin model (Neale & Cardon, 1992) uses the relative similarity of monozygotic (MZ) vs. dizygotic (DZ) twins to decompose variance in a phenotype into three latent components: variance due to additive genes (*A*), variance due to environmental influences that make twins similar to each other (the shared environment, *C*), and variance due to environmental influences that make siblings different from each other, plus measurement error (the non shared environment, *E*). These models are commonly referred to as "*ACE*" models. Based on genetic theory, the *A* components are correlated 1.0 in MZ twin pairs and 0.5 in DZ twin pairs. By definition, in all twin pairs, the *C* components are correlated 1.0, and the *E* components are uncorrelated. Variance of the ACE components are fixed to 1.0, and paths from the ACE components to each measured variable are estimated. The square of the path from the *A* component (*h*)

represents the proportion of variance in a phenotype (for example, testosterone) due to genetic influences, i.e., the heritability coefficient (h^2). Similarly, the squares of the *c* and *e* paths represent the proportions of variance due to shared environmental and non-shared environmental influences, respectively.

Our preliminary work suggests that genetic influence on testosterone differs by gender, with substantial genetic influences on testosterone in male adolescents (~60%) but minimal genetic influences on testosterone in females (Harden, Kretsch, Tackett, & Tucker-Drob, 2014). To account for these sex differences, we adjusted the twin model so that the magnitudes of the paths from the ACE components to testosterone were free to vary by gender. In addition, for opposite-sex DZ twin pairs, the correlation between the *A* component for the male twin and the A component for the female twin (rA) was freely estimated, rather than fixed to 0.5. This model is illustrated below.

Figure 2. Quantitative genetic model of testosterone in male and female twins



Note: A = Additive Genetic; C = Shared Environmental; E = Non-Shared Environmental. A, C, and E factors are standardized (M = 0, SD = 1). *m* = male; *f* = female. Model for opposite-sex DZ pairs is shown; r_A was freely estimated. For MZ pairs and same-sex DZ pairs (not shown), *r*A was fixed to equal 1.0 and 0.5, respectively.

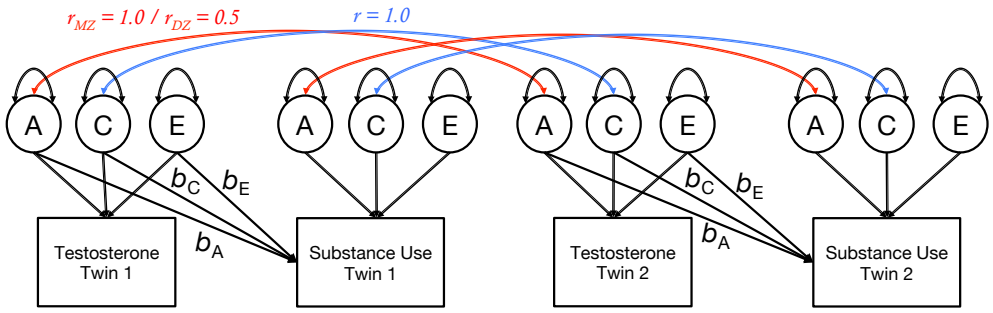
Second, we developed a measurement model of reward seeking and cognitive control using a factor analysis of self-reported impulsivity (the four subscales of the UPPS), intelligence (measured by the WASI), and performance on decision-making tasks. After examining correlations between the measures, we ran exploratory and confirmatory analyses. We expected to find at least two latent factors representing impulse control ("cold" cognition) and one reflecting reward- or sensation-seeking. Measures that did not load on latent factors were dropped from this model.

Third, we examined phenotypic associations between the factors identified in the measurement model, testosterone, and substance use. We included age, gender, self-reported pubertal timing, parent education, race, and ethnicity in all models as well. Separate models were estimated for each measure of substance use (any alcohol use, past year intoxication, and any marijuana use). These models assume a standardized normal continuous distribution underlies the observed categorical variable of alcohol use or drug use. We accounted for individuals nested within twin pairs using the "CLUSTER" option. We tested whether the specified pathways differ by gender using a multiple group model and testing whether allowing pathways to vary as a function of gender improved model fit.

Finally, we fit quantitative genetic models to any significant phenotypic associations between testosterone and substance use that were identified in the second step. The basic twin model, shown above, can be extended to include multiple phenotypes. This multivariate model allows one to decompose associations between variables into genetic, shared environmental, and nonshared environmental pathways. An

example of this extension of the twin model is shown below. In this bivariate model, substance use is regressed onto the *ACE* components of testosterone. The paths bA, bC, and bE represent associations between testosterone and substance use that are due to genetic influences, shared environmental influences, and nonshared environmental influences, respectively. A significant path bA would indicate shared genetic variance between testosterone and substance use, providing evidence that testosterone is an endophenotype.

Figure 3. Quantitative Genetic Model of Substance Use and Testosterone



Note. *A* = Additive Genetic; *C* = Shared Environmental; *E* = Non-Shared Environmental. *A*, *C*, and *E* factors are standardized (*M* = 0, *SD* = 1). *bA*, *bC,* and *bE* = common between testosterone and substance use.

RESULTS

Quantitative Genetic Models of Testosterone

Quantitative genetic models of testosterone controlled for sex-specific linear and quadratic effects of age. Standardized parameter estimates and model fit statistics for the quantitative genetic model of testosterone are shown in Table 1. For males, genetic differences accounted for approximately 50% of the variation in testosterone, shared

environmental differences accounted for approximately 20%, and the remaining variance was due to nonshared environmental differences and measurement error. For females, variance in testosterone was primarily due to shared (41%) and non-shared (42%) environmental differences. Genetic differences accounted for only 17% of variance in testosterone in females. The genetic correlation for opposite-sex twins was -0.12, $p = .88$.

Table 1. Standardized Parameter Estimates from Quantitative Genetic Model of Testosterone

	Additive Genes (h^2)	Shared Environment (c^2)	Non-Shared Environment (e^2)
Males	**0.50 (.18)***	0.20 (.27)	**0.30 (.06)***
Females	0.17 (.32)	**0.41 (.17)***	**0.42 (.07)***

$\chi^2 = 8.41$, df = 18, $p = .97$
RMSEA = 0.00
Note. Standard errors are shown in parentheses. *Significantly different than zero at $p < .05$.

Measurement Model of Reward-Seeking and Cognitive Control

Correlations between self-report measures and performance on behavioral tasks were generally low (Table 2). Self-reported sensation seeking was positively correlated with risk-taking on the BART ($r = .10$) and Stoplight Game ($r = .19$). Self-reported negative urgency was positively correlated with slope ($r = .20$) and intercept ($r = .18$) for bad decks on the Iowa Gambling Task, indicating that individuals with higher negative urgency did not learn to avoid bad decks over time. Self-reported premeditation was negatively correlated with risk-taking on the BART ($r = -.09$) and on the Stoplight Game ($r = -.10$). Intelligence was correlated with performance on several behavioral tasks, particularly with harm avoidance on the IGT (intercept $r = -.26$; slope $r = -.29$) and with fewer excess moves on the TOL ($r = -.28$).

Table 2. Correlations Between Self-report and Behavioral Measures

	1. Age	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
2. IQ	.07 n=530											
3. Premeditation	.09* n=529	.12** n=529										
4. Sensation Seeking	.11* n=529	.05 n=529	-.26*** n=529									
5. Perseverance	.09* n=529	.12** n=529	.43*** n=529	.07 n=529								
6. Urgency	.02 n=529	-.18** n=529	-.23*** n=529	.11* n=529	-.25*** n=529							
7. BART– Average pumps per balloon	.04 n=529	.10* n=529	-.09* n=522	.10* n=528	-.05 n=528	-.06 n=529						
8. TOL – Excess Moves	-.16*** n=525	-.28*** n=525	-.04 n=524	-.05 n=528	-.08 n=528	.09* n=524	-.05 n=524					
9. SLG – % Risky Decisions	.02 n=515	.02 n=515	-.10* n=514	.19*** n=514	.02 n=514	.13** n=514	.17** n=514	.02 n=510				
10. Delay Discounting	-.10* n=516	-.17** n=516	-.02 n=516	.00 n=515	-.06 n=516	.15** n=516	-.06 n=515	.14** n=511	.00 n=503			
11. IGT Good Intercept	.16** n=521	.17** n=516	.04 n=521	.04 n=521	.04 n=521	-.06 n=521	.23*** n=520	-.05 n=516	.10* n=507	-.07 n=511		
13. IGT Bad Intercept	-.06 n=521	-.26** n=521	-.04 n=521	-.03 n=521	-.05 n=521	.20*** n=521	.02 n=520	.18** n=516	.09 n=507	.17*** n=511	.08 n=521	.01 n=521

Note. *p < .05; **p < .01; ***p < .001. BART: Balloon Analogue Risk Task; IGT: Iowa Gambling Task; SLG: Stoplight Game; TOL: Tower of London. Risky decision-making on the BART was defined as the average number of times a participant pumped a balloon before "cashing out." IGT "good" and "bad" intercepts reflect performance on the final block of the task. Higher delay discounting values reflect tendency to discount delayed rewards. Sensation seeking, Perseverance, Premeditation, and Urgency are 4 subscales of the UPPS self-reported impulsivity scale.

Table 3. Twin Pair Correlations by Zygosity for Key Study Variables

	All MZ ***N*=96**	**All DZ** ***N*=165**	**MZF** ***N*=52**	**MZM** ***N*=44**	**DZF** ***N*=39**	**DZM** ***N*=50**	**DZO** ***N*=76**
Testosterone	.62** (*N*=93)	.38** (*N*=160)	.53** (*N*=51)	.70** (*N*=42)	.56** (*N*=37)	.43** (*N*=48)	.24* (*N*=75)
Pubertal Timing	.41** (*N*=96)	.12 (*N*=163)	.39** (*N*=52)	.43** (*N*=44)	.37* (*N*=39)	.17 (*N*=50)	-.02 (*N*=75)
IQ	.79** (*N*=96)	.44** (*N*=165)	.71** (*N*=52)	.85** (*N*=44)	.44** (*N*=39)	.29* (*N*=50)	.53** (*N*=76)
Lifetime Alcohol Use[1]	.65** (*N*=96)	.48** (*N*=162)	.52** (*N*=52)	.68** (*N*=44)	.47** (*N*=38)	.64** (*N*=49)	.37** (*N*=75)
Past Year Intoxication[1]	.68** (*N*=96)	.51** (*N*=163)	.64** (*N*=52)	.73** (*N*=44)	.41** (*N*=38)	.46** (*N*=49)	.61** (*N*=75)
Lifetime Marijuana Use[1]	.59** (*N*=96)	.51** (*N*=164)	.50** (*N*=52)	.69** (*N*=44)	.50** (*N*=38)	.68** (*N*=49)	.38** (*N*=75)
Sensation Seeking	.48** (*N*=96)	.19* (*N*=164)	.42** (*N*=52)	.54** (*N*=44)	.31 (*N*=39)	.15 (*N*=50)	.05 (*N*=75)
Perseveration	.45** (*N*=96)	.10 (*N*=164)	.52** (*N*=52)	.36* (*N*=44)	.14 (*N*=39)	.27 (*N*=50)	-.03 (*N*=75)
Premeditation	.40** (*N*=96)	.18* (*N*=164)	.56** (*N*=52)	.21 (*N*=44)	.22 (*N*=39)	.20 (*N*=50)	.16 (*N*=75)
Urgency	.44** (*N*=96)	.15 (*N* =161)	.51** (*N*=52)	.35* (*N*=44)	.08 (*N*=39)	.29 (*N*=50)	.10 (*N*=75)
BART– Avg pumps/balloon	.27** (*N*=96)	.10 (*N* =161)	.31* (*N*=52)	.23 (*N*=44)	.25 (*N*=39)	.15 (*N*=49)	-.01 (*N*=73)
SLG – % Risky Decisions	.20 (*N*=92)	.12 (*N* =155)	.41** (*N*=49)	-.14 (*N*=44)	.24 (*N*=39)	.13 (*N*=47)	.06 (*N*=73)
IGT Good Intercept	.29** (*N*=91)	.22** (*N*=162)	.32* (*N*=48)	.25 (*N*=43)	.12 (*N*=37)	.30 (*N*=50)	.20 (*N*=75)
IGT Bad Intercept	.35** (*N*=91)	.09 (*N*=162)	.37* (*N*=48)	.34* (*N*=43)	.14 (*N*=37)	.25 (*N*=50)	-.04 (*N*=75)
Delay Discounting	.42** (*N*=91)	.24** (*N*=158)	.39** (*N*=49)	.46** (*N*=42)	.28 (*N*=38)	.35* (*N*=49)	.04 (*N*=71)
TOL – Excess Moves	.31** (*N*=93)	.22** (*N*=163)	.25 (*N*=50)	.43** (*N*=43)	.28 (*N*=39)	.29* (*N*=49)	.06 (*N*=75)

Note. *p < .05; **p < .01. [1]Tetrachoric correlations for dichotomous variables. MZM= male monozygotic twins; MZF=female monozygotic twins; DZM=male dizygotic twins; DZF=female dizygotic twins; DZO=opposite sex dizygotic twins.

Twin correlations for all study measures are shown in Table 3. Consistent with results from the quantitative genetic model of testosterone, in males, MZ correlations were substantially higher than DZ correlations, whereas in females, twin correlations did not differ by zygosity. Correlations for intelligence, and for the behavioral tasks that loaded

on the cognitive ability factor, were higher among MZ twins than among DZ twins, indicating substantial genetic influence on this phenotype. Tetrachoric correlations for the three substance use measures were all significant but did not differ by zygosity. These results suggest variance in substance use in this sample was primarily due to the shared environmental (i.e., family level) differences between twins and not due to genetic differences.

Based on previous exploratory analyses (Kretsch et al., 2014), we ran a confirmatory factor analysis of three factors to identify latent constructs underlying phenotypic correlations. We specified a CFA model in which all indicators were standardized and regressed on age and gender. This model fit the data well ($\chi^2 = 43.941$, $df = 36$, $p = .17$, RMSEA = .02, CFI = .977, TLI = .957). Factor loadings, factor correlations, and regression coefficients for this model are summarized in Table 4. Factor 1 reflected general *Cognitive Ability*. IQ loaded on this factor, as well as performance on the TOL. Both measures of performance on the IGT—avoidance of bad decks and preference for good decks by the end of the task—loaded on Factor 1. Urgency loaded negatively on this factor. Preference for immediate reward on the Delay Discounting task also loaded negatively on this factor. Factor 2 reflected self-reported *Impulsivity*. Three of the four subscales on the UPPS (premeditation, perseverance, and urgency, but not sensation seeking) loaded on this factor. Factor 3 seemed to reflect overall "approach" behavior—including risk-taking on the SLG and the BART, as well as self-reported sensation-seeking. Self-reported perseverance also loaded positively on

this factor. This factor represented the closest conceptual overlap with reward seeking and was therefore expected to correlate with testosterone. We refer to Factor 3 as *Reward Seeking*.

Table 4. Standardized Parameter Estimates from CFA

	F1 Cognitive Ability	F2 Impulsivity	F3 Reward Seeking
Self report measures	***Factor Loadings***		
UPPS Premeditation		-.67***	
UPPS Sensation Seeking			0.63***
UPPS Perseverance		-1.10***	0.76**
UPPS Urgency	-0.31***	0.31***	
Behavioral measures			
WASI – IQ	0.62***		
BART– Average pumps per balloon	0.11*		0.14*
TOL – Excess moves	-0.43***		
SLG – Percent risky decisions			0.29***
Delay Discounting	-0.32***		
IGT Good Intercept	0.19**		
IGT Bad Intercept	0.44***		
	Residual Correlations		
BART and SLG	.12**		
IGT Good and IGT Bad Intercepts	.17***		
IGT Good Intercept and BART	.19***		
	Factor Correlations		
F1 and F2	0.14		
F2 and F3	0.62***		
F1 and F3	0.02		

Phenotypic Associations between Testosterone, Reward seeking, and Substance Use Initiation

In the next set of analyses, the three factors (Cognitive Ability, Impulsivity, and Reward Seeking) were regressed on testosterone, self-reported pubertal timing, parental education, black race, and Hispanic ethnicity. We regressed each indicator of the factors

on age and gender. Separate models were run for each dichotomous measure of substance use (any alcohol use, past year intoxication, and any marijuana use). Substance use was also regressed onto testosterone, pubertal timing, parent education, race, and ethnicity. Of primary interest were the associations between testosterone, reward seeking, testosterone and substance use, and reward seeking and substance use. Results from phenotypic analyses are shown in Table 5 and are illustrated in Figure 4.

Contrary to predictions, there was no association between testosterone and any measure of substance use and no association between testosterone and reward seeking. Testosterone did not predict greater likelihood of alcohol use (OR = 1.04, p = .58), past year intoxication (OR = 1.03, p = .68), or marijuana use (OR = .82, p = .62), and it was not associated with reward seeking (β= -.04, p = .57).

We also tested whether reward seeking and the other latent factors predicted substance use initiation. Surprisingly, the effects of reward seeking on substance use variables were not significant, although the effect on any alcohol use was in the expected direction (OR = 1.37, p = .06). The other two latent factors, however, did predict substance use. Adolescents with higher cognitive ability (F1) were less likely to report alcohol use (OR = .71, p < .01), past year intoxication (OR = .74, p < .05), and ever smoking marijuana (OR = .54, p < .001). More impulsive adolescents (F2) were more likely to report ever drinking (OR = 1.37, p < .01) and past year intoxication (OR = 1.35, p = .05).

Additional models (not shown in Table 5) added interactions between reward seeking and the two other latent factors as predictors of substance use. In the model

predicting alcohol use, there were no significant interactions between reward seeking and cognitive ability or reward seeking and impulsivity. However, when the interaction between reward seeking and cognitive ability was added to the model (p-value?), the main effect of reward seeking on alcohol use was statistically significant (OR = 1.73, $p <$ .05). This interaction suggested that the effects of reward seeking on alcohol use were stronger among adolescents with lower cognitive ability.

Table 5. Standardized Parameter Estimates for Models Predicting Substance Use Initiation

	Any alcohol	Past year intoxication	Any marijuana
Odds Ratio (SE)			
Substance use on F1	0.71 (.29)**	0.74 (.36)**	0.54 (.25)***
Substance use on F2	1.37 (.58)**	1.35 (.63)*	0.97 (.54)
Substance use on F3	1.23 (.54)	1.23 (.60)	0.89 (.54)
Substance use on Age	1.58 (.35)***	1.58 (.38)***	1.18 (.29)***
Substance use on SES	1.01 (.33)	0.94 (.32)	0.84 (.29)
Substance use on Black	0.61 (.34)***	0.56 (.30)***	0.48 (.29)***
Substance use on Hispanic	0.73 (.62)	0.80 (.68)	0.55 (.72)
Substance use on PDS	1.04 (.29)	1.02 (.35)	0.78 (.26)
Substance use on T	1.04 (.08)	1.03 (.33)	0.82 (.26)
Regression coefficients (SE)			
F1 on SES	0.35 (.09)***	0.34 (.09)***	0.35 (.09)***
F1 on Black	-1.06 (.24)***	-0.97 (.23)***	-1.06 (.24)
F1 on Hispanic	-0.73 (.20)***	-0.71 (.19)***	-0.72 (.20)***
F1 PDS	0.18 (.07) **	0.17 (.06)**	0.17 (.07)**
F1 on T	-0.04 (.07)	-0.04 (.07)	-0.03 (.07)
F2 on SES	-0.13 (.07)	-0.13 (.06)*	-0.12 (.07)
F2 on Black	-0.22 (.21)	-0.27 (.20)	-0.21 (.21)
F2 on Hispanic	-0.33 (.17)	-0.33 (.06)*	-0.31 (.18)
F2 on PDS	-0.03 (.06)	-0.02 (.07)	-0.04 (.07)
F2 on T	0.02 (.07)	0.01 (.07)	0.02 (.07)
F3 on SES	-0.15 (.06)*	-0.16 (.06)*	-0.11 (.06)
F3 on Black	-0.27 (.21)	-0.33 (.22)	-0.29 (.20)
F3 on Hispanic	-0.39 (.17)*	-0.40 (.17)*	-0.38 (.17)*
F3 on PDS	0.04 (.06)	0.04 (.06)	-0.11 (.09)

Table 5, cont.

F3 on T	-0.01 (.07)	-0.04 (.07)	-0.04 (.06)
	-0.15 (.20) /	-0.16 (.20) /	-0.16 (.20) /
PDS on Black[a]	-0.28 (.18)	-0.28 (.18)	0.14 (.18)
	0.16 (16) /	0.16 (.20) /	0.16 (.20) /
PDS on Hispanic[a]	-0.68 (.14)***	-0.68 (.14)***	-0.68 (.14)***
	0.14 (.22) /	0.14 (.22) /	0.14 (.22) /
T on Black[a]	0.28 (.19)	0.28 (.19)	0.28 (.19)
	0.27 (.21)/	0.28 (.21) /	0.28 (.21) /
T on Hispanic[a]	0.09 (.24)	0.01 (.22)	0.09 (.24)
SES on Black	-0.50 (.12)***	-0.58 (.14)***	-0.50 (.12)***
SES on Hispanic	-1.04 (.11)***	-1.04 (.11)***	-1.04 (.11)***
Correlations			
	-0.01 (.07) /	-0.01 (.07) /	-0.01 (.07) /
PDS with SES[a]	0.32 (.06)***	0.26 (.05)***	0.32 (.06)***
	-0.05 (.06) /	0.06 (.07) /	-0.05 (.06) /
T with SES[a]	0.07 (.08)	0.07 (.08)	0.07 (.08)
	0.25 (.07)*** /	0.25 (.07)*** /	0.24 (.07)*** /
T with PDS[a]	-0.01 (.05)	-0.01 (.05)	-0.03 (.05)
Thresholds			
	0.59 (.13)*** /	0.89 (.16)*** /	0.85 (.15)*** /
Substance use[a]	0.83 (.14)***	1.11 (.17)***	1.13 (.14)***
Model fit indices			
χ^2	243.10	239.25	229.18
df	229	229	229
p	.25	.31	.48
RMSEA	.02	.01	.02
CFI	.98	.98	1.00
TLI	.97	.98	1.00

*$p < .05$; **$p < .01$; ***$p < .001$

[a]Parameter estimates for males, followed by estimates for females. All other estimates shown did not differ between males and females.

F1 = cognitive ability; F2 = impulsivity; F3 = reward seeking; T = Testosterone, standardized by age, age^2 and sex, log transformed and winsorized; PDS = Pubertal Development Scale, standardized by age and sex; SES = parent education in years.

Gender Differences in Phenotypic Associations

We examined gender as a moderator of the paths between testosterone, substance use, and reward seeking using a series of multiple group nested model comparisons. The baseline (most constrained) model constrained all parameters to be equal across gender except for substance use thresholds, means of factor indicators, and correlations between

testosterone and pubertal timing. Freeing the path between testosterone and reward seeking across gender did not significantly improve model fit in the models for alcohol use ($\Delta\chi^2 = .75$, $\Delta df = 1$, $p = .39$), intoxication ($\Delta\chi^2 = .87$, $\Delta df = 1$, $p = .35$), or marijuana use ($\Delta\chi^2 = .69$, $\Delta df = 1$, $p = .41$). Model fit did not improve when the direct path between testosterone and alcohol use was free to vary between genders ($\Delta\chi^2 = .55$, $\Delta df = 1$, $p = .46$) or when the path between reward seeking and alcohol use was free to vary between genders ($\Delta\chi^2 = .48$, $\Delta df = 1$, $p = .49$). Similarly, model fit did not improve when the paths between testosterone and intoxication or ($\Delta\chi^2 = .93$, $\Delta df = 1$, $p = .34$) between reward seeking and intoxication ($\Delta\chi^2 = .01$, $\Delta df = 1$, $p = .92$) between gender. Finally, model fit did not improve when the paths between testosterone and marijuana use was free to vary between gender ($\Delta\chi^2 = .36$, $\Delta df = 1$, $p = .55$) or when the path between reward seeking and marijuana use were free to vary between gender ($\Delta\chi^2 = .67$, $\Delta df = 1$, $p = .42$) . Thus, gender did not appear to moderate associations between testosterone and reward seeking or between testosterone and any measure of substance use.

Bivariate Quantitative Genetic Models

The final step in the planned analyses was to decompose significant phenotypic associations between testosterone, alcohol use, and reward seeking into genetic, shared environmental, and non-shared environmental pathways. As shown in Figure 3, a significant pathway, bA, between testosterone and an observed phenotype would suggest that they were correlated due to common genetic variance. This would provide evidence that testosterone was an endophenotype (Gottesman & Gould, 2003). Surprisingly, we found no significant phenotypic associations between testosterone and substance use or

reward seeking. This precluded further inquiry into testosterone as an endophenotype. Therefore, no bivariate quantitative genetic models of testosterone, reward seeking, and substance use were estimated.

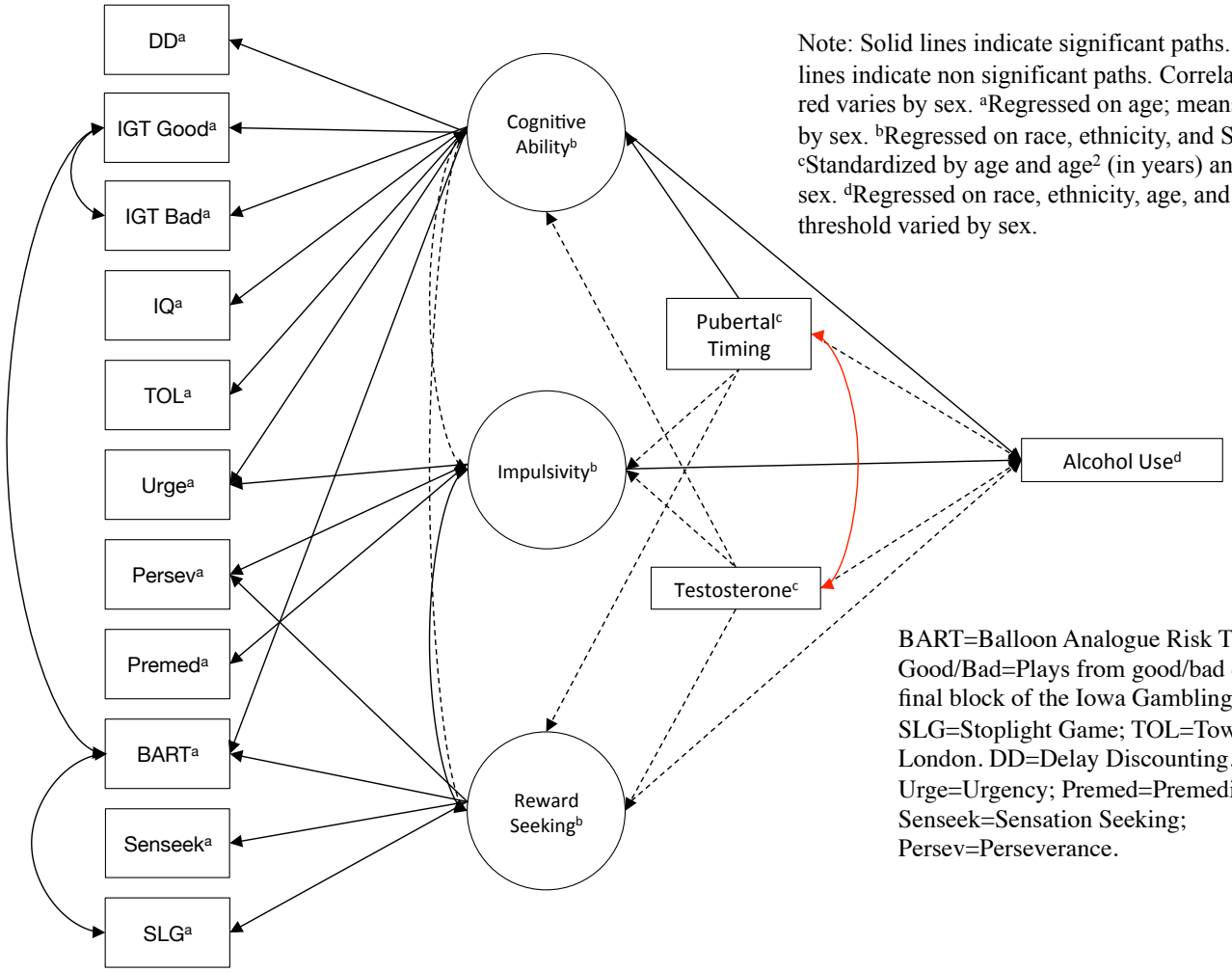


Figure 4. Full Model Predicting Alcohol Use

Note: Solid lines indicate significant paths. Dashed lines indicate non significant paths. Correlation in red varies by sex. [a]Regressed on age; means varied by sex. [b]Regressed on race, ethnicity, and SES. [c]Standardized by age and age^2 (in years) and by sex. [d]Regressed on race, ethnicity, age, and SES; threshold varied by sex.

BART=Balloon Analogue Risk Task; IGT Good/Bad=Plays from good/bad decks on final block of the Iowa Gambling Task; SLG=Stoplight Game; TOL=Tower of London. DD=Delay Discounting. Urge=Urgency; Premed=Premediation; Senseek=Sensation Seeking; Persev=Perseverance.

DISCUSSION

The current study examined testosterone as a potential endophenotype for substance use in adolescents, using a sample of high school aged twins. It was hypothesized that salivary testosterone would predict initiation substance use via its effects on reward-seeking or sensation-seeking, that this association would be stronger for males than for females, and that this association would be due to common genetic effects. Reward seeking was measured as a latent construct using a battery of self-report and behavioral tasks. The assessment battery revealed three latent factors, including reward seeking, which predicted initiation of alcohol use in both males and females. Contrary to predictions, there were no indirect or direct effects of testosterone on initiation of substance use in either males or females.

The lack of associations between testosterone in substance use in females is consistent with previous studies (e.g., de Water et al., 2013; Martin et al., 1999). However, these findings are inconsistent with the extant body of research showing associations between testosterone, risk-taking, and substance use in males (de Water et al., 2013; Costello et al., 2007; Erikkson et al., 2005). Moreover, results do not support testosterone as an endophenotype for risk-taking in adolescence. Testosterone levels in this sample were highly heritable in males, but there was no phenotypic pathway between testosterone and reward seeking or substance use. Many factors may have contributed to these null findings, and below, we discuss five likely explanations.

First, the level of substance use in this sample was low compared with national averages. We examined cross-sectional relationships between testosterone and alcohol

and drug use initiation in a sample that was predominantly abstinent from substance use. Alcohol use was uncommon – less than a quarter of the current high school aged sample reported ever using alcohol. In contrast, the prevalence of lifetime alcohol use among 10th graders in the most recent wave of the Monitoring the Future study (Johnston et al., 2014) was 49.3%. Past year intoxication in our sample was 17%, compared with 24.6% of 10th graders surveyed in Monitoring the Future. Marijuana use in the current sample (18%) was also lower than national average in MTF; this trend is consistent with overall lower rates of marijuana use in the state of Texas (SAMHSA, 2014). A weak association between testosterone and substance use would be difficult to detect in this study sample.

Second, the age range and cross-sectional design of the current study sample may explain findings regarding testosterone and substance use. Some research suggests that testosterone may predict substance use longitudinally, with effects not emerging until young adulthood, when there is more variability in levels of substance use. For example, a longitudinal study of adolescent boys found serum testosterone levels in early adolescence (ages 12-14) predicted social potency (i.e., interpersonal forcefulness and leadership) and approval of aggressive/antisocial behavior at age 16, which, in turn, predicted illicit substance use at age 19 (Reynolds et al., 2007). This sample was also at elevated genetic risk for substance use problems due to positive family history in 40% of the sample. Another longitudinal study found testosterone mediated the association between neighborhood-level socioeconomic adversity and aggressive behavior, which, in turn, predicted cannabis use in young adulthood (Tarter et al., 2009). In these studies,

testosterone measured in early adolescence has predicted intermediary outcomes that subsequently influenced substance use later in the lifespan.

Third, the phenotypes we analyzed in the current study were mild forms of substance use: initiation of alcohol and marijuana use and a dichotomous measure of past year alcohol intoxication. Some research has found stronger associations between testosterone and problematic or heavy substance use than between testosterone and substance use initiation (Eriksson et al., 2005). We had predicted that the effect of testosterone would be mediated by reward seeking. If testosterone affects the response to the rewarding properties of alcohol and/or marijuana, this effect would not appear until after an individual has tried these substances. Substance use initiation may be more influenced by environmental factors than by genetic factors (such as genetic variation in testosterone). Twin studies have shown that initiation of substance use is determined primarily by shared (familial) environmental factors, whereas established patterns of substance use and substance use disorders are more influenced by genes (Rose & Dick, 2004; Hopfer et al., 2003; Han, McGue, & Iacono, 1999; Huizink et al., 2010). The strongest predictors of substance use initiation in the current study were age, race, and cognitive ability—all of which may be proxies/indicators of shared environmental effects. Thus, the results of the current study, which assessed for substance use initiation only, may reflect differential associations between genes, testosterone and substance use across levels of substance use severity. Testosterone may play a role in whether an adolescent who has tried alcohol progresses to heavier use in young adulthood.

Fourth, the absence of an association between testosterone and reward seeking and substance use may be due to interactions between testosterone and other variables. One unexplored possibility is that testosterone is a moderator, not a mediator, of genetic influence on risk-taking. There is emerging evidence that gonadal hormones moderate genetic influences on other behaviors in adolescence. For example, recent twin research suggests that estradiol amplifies genetic influence on disordered eating behaviors in adolescent girls (Klump et al., 2012). In a study of adult men, testosterone interacted with the MAOA genotype to influence antisocial behavior (Sjoberg et al., 2008). Although, to our knowledge, no studies have examined testosterone as a moderator of genetic influence on substance use, the androgen receptors distributed throughout the nervous system that bind with testosterone are themselves transcription factors that regulate gene expression (Witt, 2008; Nilsson & Gustafsson, 2000). This possibility that testosterone moderates genetic influence on substance use is a plausible alternative to the endophenotype hypothesis that we tested in this study. If testosterone affects reward seeking and/or substance use by activating genetic vulnerabilities, its effects may be apparent only among individuals at high genetic risk.

In addition to possible gene-by-hormone interactions, testosterone may predict risk-taking and substance use through interactions with other hormones that were not included the current study. In particular, there is growing evidence that testosterone and cortisol interact to influence risk-taking and related behaviors. According to the *dual hormone hypothesis*, testosterone is positively related to status-related behaviors only when cortisol levels are low and not when they are high (Mehta & Josephs, 2010; Pompa

et al., 2007; Dabbs et al., 1991). This interaction has been found in studies of aggression, dominance, social status, and externalizing psychopathology (e.g., Pompa et al. 2007; Edwards & Casto, 2013; van Den Bos et al., 2013; Tackett et al., 2014), as well as self-reported, informant-reported, and behavioral measures of risk-taking (Mehta et al., 2015). Cortisol is thought to moderate testosterone's influence on behavior via joint effects on the HPA and HPG axes (Viau, 2002), through down-regulation of the mesolimbic reward pathway (Montoya et al., 2014), or through inhibition of orbitofrontal cortex regions involved in impulse control (Peters et al., 2015). Risk-taking in general, and substance use in particular, are strongly tied to social status in adolescence. The dual hormone hypothesis might explain the lack of direct and indirect effects of testosterone on substance use initiation in the current study. Sex-specific interactions between gonadal and stress hormones contribute to sex differences in alcohol use that emerge during puberty late adolescence (reviewed by Witt, 2008). However, the proposed mechanism for why testosterone and cortisol interact and lead to sex differences in alcohol use is through an effect on response to alcohol (Witt, 2008). Thus, an interaction between testosterone and cortisol may be less relevant in predicting alcohol and drug use initiation and more relevant for predicting progression of substance use following initiation.

Limitations

These findings should be interpreted in light of several limitations. Only one testosterone sample was taken from each participant. Although we accounted for variability in time of day in our data, there are individual differences in how hormone levels change throughout the day, which we did not measure. We relied on females' self-

report of menstrual cycle phase, which may not have been completely reliable. As noted above, the age range and cross sectional design of the current study are limitations. Females in the study were nearly all post-menarcheal, and all adolescents were at least mid-pubertal. Individual differences in testosterone may be more predictive of risk behavior—inside and outside of the laboratory—at different stages of development; this has been shown for the testosterone-aggression association, which differs in adults vs. adolescents (Archer, 2006). Several studies that have found large associations between testosterone and behavioral measures have included more age-heterogeneous samples and/or younger or older samples. For example, Peper, Koolschijn, and Crone (2013) reported significant associations between testosterone and risk-taking on the BART from a study sample that ranged in age from 8 to 25. In that study, the correlations between testosterone, pubertal development, and age were significant in both males and females and strikingly larger than those found in the current study sample. De Macks et al. (2011) found that testosterone in 10-16 year olds correlated with activity in the ventral striatum in response to receiving monetary rewards in a gambling task. Other studies linking testosterone to performance on the IGT have used adult samples (e.g., van Honk et al., 2004; Stanton et al., 2011; Reavis & Overman, 2001). Longitudinal studies (e.g. Braams et al., 2015) can elucidate hormonal effects on both within-person and between-person variability in risk-taking.

Conclusion

Animal and human research has revealed associations between testosterone and forms of risk-taking including substance use. The effect of testosterone on reward

circuitry in the brain may mediate this association by sensitizing adolescents to the rewarding properties of risk-taking. Genetic variation in testosterone has been hypothesized to account for the genetic influence on substance use that emerges in adolescence. The current study used a twin design and examined testosterone as a potential endophenotype for initiation of alcohol use and marijuana use. After controlling for age, gender, race and ethnicity, there were no associations between testosterone and substance use or between testosterone and reward seeking. The contrast between current study findings and previous research may be due to a number of factors, and determining the best explanation will require future research and dissemination of both expected and unexpected results.

Chapter 3: Pubertal Timing and Peer Influence on Risk-Taking[1]

BACKGROUND

The increasing salience of peer relationships is a hallmark of adolescent social development. Compared to children and adults, adolescents are more sensitive to the influence of their peers (Gardner & Steinberg, 2005; Chein et al., 2010; Steinberg & Monahan, 2007). Although the growing importance of peers reflects a normal developmental process, peer influence is often studied in the context of maladaptive behaviors, such as substance use, delinquency, and other forms of risk-taking. Adolescents are more likely to engage in these behaviors in the presence of peers (Chassin, Hussong, & Beltran, 2004; Zimring, 1998; Ouimet et al., 2010), and longitudinal and experimental studies provide convergent evidence for a causal association between exposure to peer risk-taking and individual risk-taking (Jaccard, Blanton, & Dodge, 2005; Cruz, Emery, & Turkheimer, 2012; Gardner & Steinberg, 2005). Intervention and prevention efforts have attempted to understand peer influence processes both in terms of preventing negative influence (e.g., iatrogenic effects of group treatment for adolescents, reviewed by Dishion & Dodge, 2005) and garnering positive influence (e.g., "Above the Influence" anti-drug campaign; Slater et al., 2011).

Of particular interest are factors that render certain adolescents more vulnerable to negative peer influence than others (see Brechwald & Prinstein, 2011, for a review). Pubertal timing may be one factor that contributes to individual differences in susceptibility to peer influence. Among adolescent girls, early pubertal timing predicts an array of negative psychosocial outcomes, including heightened involvement in risky behavior (Mendle, Turkheimer, & Emery, 2007; Buchanan, Eccles, & Becker, 1992;

[1] Kretsch, N., Mendle, J. & Harden, K. P. (in press). A Twin Study of Objective and Subjective Pubertal Timing and Peer Influence on Risk-Taking. *Journal of Research on Adolescence*.
The first author conceived of the study, performed statistical analyses, and drafted the manuscript.

Stattin, Kerr, & Skoog, 2011). Moreover, behavioral genetic research suggests that pubertal timing moderates the relative influences of genes and environment on risk-taking; environmental influences on delinquency are stronger (and genetic influences weaker) for girls with early pubertal timing (Burt, McGue, DeMarte, Krueger, & Iacono, 2006; Harden & Mendle, 2012). One of these environmental influences may be the peer group, as several previous studies have found interactions between peer risk behavior and pubertal timing in predicting delinquency (Caspi, Lynam, Moffitt, & Silva, 1993; Fergusson et al., 2007; Stattin et al., 2011) and substance use (Costello et al., 2004; Biehl, Natsuaki, & Ge, 2007).

The importance of the peer group is highlighted in nearly all of the theories proposed to explain the negative correlates of early pubertal timing. For example, the *maturation disparity hypothesis* (reviewed in Ge & Natsuaki, 2009) focuses on the gap between physical and psychological development that is widened for early maturing adolescents. Pubertal development brings substantial social changes, including upheaval in long-standing friendships and attention from opposite-sex peers. Adolescents who face these transitions at younger ages are less psychologically mature and may lack the cognitive and emotional resources necessary to cope with them (Ge & Natsuaki, 2009). The *contextual amplification hypothesis* (Ge, Brody, Conger, Simons, & Murry, 2002) considers pubertal development in the context of adolescents' social background. The negative effects of pubertal timing may be intensified in adverse environments, where girls must confront numerous risk factors, including greater exposure to delinquent peers. In support of this, the association between early maturation and deviant behavior is attenuated for girls in same-sex schools, presumably because they are "protected" from the influence of boys (Caspi et al., 1993). Other variants, such as the *peer socialization hypothesis* (Stattin et al., 2011) suggest that early maturation motivates affiliation with

deviant peers because early-maturing adolescents select friends who are similar to them; these friends are likely to be older and to engage in more norm-breaking behavior than same-age, typically developing peers. This motivation for an alternate social group may be fueled by difficulty maintaining friendships with same-age peers, as early maturing adolescents tend to experience rejection and even victimization from other children their age (Haynie & Piquero, 2006; Reynolds & Juvonen, 2011). Finally, affiliation with deviant peer groups may represent a reaction to the fact that adolescents are reproductively mature but are still not considered adults (Moffitt, 1993). This theory of adolescent-limited delinquency may explain why early maturers, who experience an extended period between reproductive and sociocultural maturity, are at heightened risk for antisocial behavior.

Measurement of Puberty

Much of the research reviewed above has used age at menarche as a measure of pubertal timing. Menarche is a discrete and personally salient event that is recalled with moderate-to-high reliability (Casey et al., 1991; Dorn, Sontag-Padilla, Pabst, Tissot, & Susman, 2013). In considering the intersection of puberty and peer environment, different indicators of pubertal timing may be more or less relevant. For example, menarche is a private event, whereas breast development is a visible social signal of reproductive maturity and potential readiness to engage in romantic or sexual relationships (Petersen & Taylor, 1980; Brooks-Gunn & Warren, 1989; Blyth, Simmons, & Zakin, 1985). There is growing recognition that different measures of pubertal development may differentially predict behavior (e.g., Carter, Silverman, & Jaccard, 2013; Michael & Eccles, 2003; Natsuaki et al., 2009). In particular, "subjective" measures of pubertal development, which rely on girls' self-reports of how they perceive their own bodies, have been

criticized for their poor correspondence with objective benchmarks like age at menarche and with physician ratings of physical development (Dorn et al., 2006). Yet these subjective self-perceptions may nevertheless be psychologically meaningful. Even if a girl is objectively "on time" (i.e., her development is typical of her age group), if she perceives herself as more developed, and if she judges these physical changes negatively, she may be at risk for psychological problems (Brooks-Gunn & Warren, 1989; Brooks-Gunn, Attie, Burrow, Rosso, & Warren, 1989; Rierdan et al., 1988). In the case of risk-taking, if delinquent peer affiliation and delinquent behavior are attempts to assume adult status before one is granted adult privileges by society (Moffit, 1993), then whether an adolescent "feels like" an adult may matter more than a physician's assessment of her Tanner stage. In fact, subjective measures of pubertal development are often more potent predictors of health outcomes than more objective measures (e.g., Graber et al., 1997, 2004; Lanza & Collins, 2002; Wichstrom, 2001; Deppen et al., 2012; Stice et al., 2001).

In addition, subjective perceptions of pubertal timing may be particularly relevant for understanding effects of puberty that persist into later adolescence. By middle adolescence, when rates of risk-taking increase (FBI, 2012; CDC, 2009), most girls have experienced menarche and have "caught up" to each other in terms of pubertal stage. Yet older adolescents may continue to *perceive* themselves as early maturers or late bloomers. Consistent with this idea, Cance, Ennett, Morgan-Lopez, and Foshee (2012) analyzed longitudinal data on perceived pubertal timing from a large sample of adolescents aged 11-17, and concluded that "while pubertal development is a dynamic process, perceptions of pubertal timing based on early adolescent experiences are stable throughout adolescence and contribute to adolescent identity development" (p. 775). Thus, subjective measures of pubertal timing in older adolescents may be tapping

psychologically meaningful differences in identity that are rooted in the pubertal transition.

There are important distinctions between types of self-report measures of pubertal timing. Some measures, which have been referred to as "stage-normative" pubertal timing (Cance et al., 2012) ask adolescents to rate aspects of physical development (e.g., breast growth, body hair, height, voice changes) on a scale with visual or descriptive anchors, and these ratings are normed by age. Several studies have identified interactions between the peer environment and pubertal timing using these stage-normative measures (Costello et al., 2007; Biehl et al., 2007; Fergusson et al., 2007). Other measures assess "peer-normative" pubertal timing, and ask adolescents to rate their development compared to their same-age peers. Peer comparison measures, while correlated with stage-normative measures, are also unique predictors of health outcomes (Yuan, 2007; Carter et al., 2013; Harden, Mendle, & Kretsch, 2011). Moreover, because these measures provide insight into an adolescent's self-image in relation to her peer group, they are particularly relevant to understanding peer influence, the focus of the current study.

Measuring Peer Influence

Studies of peer influence vary in the way that peer influence is defined and measured, for both empirical and theoretical reasons. Prevailing theories of peer influence all recognize the interplay between selection and socialization. Understanding when and how peers influence individual risk-taking requires controlling for the fact that adolescents tend to affiliate with peers who share behavioral characteristics. A twin design is a "quasi-experimental" research design that is useful in parsing selection versus socialization effects. This approach tests whether twins who are discordant for exposure

to an environmental variable, such as peer risk-taking, are also discordant for a behavioral outcome (Lahey & D'Onofrio, 2010; Heath et al., 1993; Silberg et al., 2003). Previous twin and sibling studies of peer influence have generally supported a socialization model of peer influence, while also identifying selection effects. In a study of siblings and their best friends, Harden et al. (2008) found genetic influences on an individual's substance use also influenced his or her exposure to peer substance use (a gene-environment correlation). At the same time, exposure to peer substance use predicted individual substance use after controlling for gene-environment correlation. Studies that have used the twin design in conjunction with longitudinal and social network analysis (e.g., Cruz et al., 2012) have found further support for a causal model of peer influence on risky behavior. The current study employs the twin design to test whether the "quasi-causal" association between peer and individual risk-taking—that is, the within-twin pair association that remains after controlling for between-family differences in genetic and environmental background factors—is moderated by pubertal timing.

"Peer influence" is frequently inferred from the similarity between peer behavior and individual behavior; making this inference often requires that adolescents report on the behavior of their peers (Costello et al., 2007; Dick, Rose, Viken, & Kaprio, 2000). Perception of peer behavior is an important construct in itself, as perceived norms are one of the strongest correlates of adolescents' willingness to engage in risky behavior (Neighbors et al., 2007). However, adolescents also tend to overestimate the similarity to their peer group (Jussim & Osgood, 1989), which may lead researchers to overestimate peer influence (Gottfredson & Hirschi, 1991; Yun, Cheong, & Walsh, 2011). One approach that overcomes this potential limitation is the use of a direct report of peer risk-taking. Such data are available from studies in which adolescents nominate friends from

school rosters, and the delinquency reported by one's nominated friends is used as a measure of exposure to risky peers. Following previous research on pubertal timing and peer influence on delinquency (Fergusson et al., 2007), the current study used peer-reported behavior as a more conservative measure of peer group similarity.

Goals of the Current Study

We addressed two research questions. First, after controlling for genetic and environmental selection factors, does affiliation with risk-taking peers predict individual risk-taking? Based on previous research, particularly genetically-informed studies, we expected that affiliation with risk-taking peers would predict individual risk-taking after controlling for selection factors. Second, is the relationship between peer and individual risk-taking moderated by pubertal timing? We predicted that girls with earlier pubertal timing would be more susceptible to peer influence on risk-taking. That is, the relationship between peer and individual risk-taking that remained after controlling for selection factors would be strongest for early maturing girls. We used three indicators of pubertal timing: age at menarche, self-rated body changes, and a peer comparison item that asked girls to compare their own physical development to other girls of the same age. As reviewed above, previous studies have found interactions between pubertal timing and peer context using both objective and subjective measures. Thus, we predicted moderation by all of these measures.

METHOD

Participants

The current study used data from the National Longitudinal Study of Adolescent Health (Add Health; Udry, 2003), a longitudinal study of health and risk behaviors among adolescents in the United States. Add Health used a school-based sampling

procedure, in which a survey was administered to all students in participating schools. Schools in the United States with at least 30 enrollees were stratified by geographic region, demographic composition, and school type, and a random sample of schools was selected from these strata. Of these schools, 134 (79%) agreed to participate, yielding a sample of 90,118 adolescents who participated in the initial, in-school survey, administered in 1994-95. The school survey included items that queried whether the adolescent had a twin. An in depth, in-home interview was conducted with a subsample of 20,745 adolescents (10,480 females) who were ages 12-19 at Wave I (1994-95). Since this initial interview, three follow up interviews have been conducted, in 1995-96 (Wave II), 2001-02 (Wave III) and 2007-08 (Wave IV).

The Add Health study deliberately oversampled sibling pairs for the in-home interview. A total of 253 female-female twin pairs completed the Wave I home interview. Twin zygosity was assessed using 11 molecular genetic markers (Smolen & Hewitt, 2003), by self-reports, and by responses to four questions concerning similarity of physical appearance. These questionnaire items have been cross-validated by genetic analyses (Spitz et al., 1996). A previous analysis of the Add Health twin sample indicated that demographic characteristics of the twin sample did not differ from the full sample (Jacobson & Rowe, 1999). The current study sample included 248 female twin pairs (496 individuals; 137 monozygotic pairs, 111 dizygotic pairs), ages 12 -19 (M=16.0, SD=1.5) who participated in the in-home interview. Five pairs were excluded because they did not have any data on pubertal timing, individual risk-taking, or peer risk-taking. Race and ethnicity, reported by adolescents, was classified as White (53.2%), African American (24.4%), Hispanic (15.6%), or Other, including Asian and Native American (6.8%).

Measures

Risk-taking

Risk-taking was measured using seven items from the in-school survey. Adolescents were asked how often in the past year they had smoked cigarettes, consumed alcohol, gotten drunk, raced on a bike or skateboard or in a car or boat, lied to their parents, skipped school, and done something dangerous because they were dared to (α = .73). Frequency was based on a six-point scale: *never* (0), *one or two times* (1), *once a month or less* (2), *2-3 days a month* (3), *once a week* (4), *3-5 days a week* (5), and *nearly every day* (6). Scores were averaged to obtain a mean level of risk-taking for each adolescent.

Peer risk-taking

Peer risk-taking was assessed using friendship nominations to identify each adolescent's friends. On the in-school survey, adolescents nominated up to five male and five female friends. From these nominations, we calculated the average risk-taking (based on the same seven items used to calculate individual risk-taking) reported by peers who either nominated or were nominated by the "target" adolescent. Limitations of the peer risk-taking data should be noted. Adolescents were allowed to nominate friends who did not attend the same school but data were only collected from identifiable in-school nominations. Nominations were non-identifiable if the nominee was not on the school roster or not in the study. Of the 496 individuals in the current sample, 139 (28%) had no peer data because they had no identifiable peer nominations. These individuals were more likely to be Hispanic ($\chi^2 = 7.87$, $df = 1$, $p < .01$). There were no other demographic differences and no differences in average risk-taking between adolescents with and without identifiable peer nominations. These individuals were retained in the sample

because they were informative regarding covariation between twins' individual risk-taking scores.

Other peer characteristics

Although not all nominations were identifiable, all were classified as either male or female. For adolescents who nominated at least one friend, the proportion of nominated friends who were male and the average age of nominated friends were calculated.

Pubertal timing

Three measures of pubertal timing were used in the current analyses.

Age at menarche. Participants reported at Waves I and II if they had "ever had a menstrual period" and, if so, during which month and year they had experienced their first menstrual cycle. At Wave III, they were asked "how old were you when you got your period for the first time?' The current study used the earliest reported age at menarche for each adolescent (i.e., if an adolescent initially reported that she began menstruating at age 12 and subsequently reported that she began menstruating at age 13, age 12 was used as age at menarche). The first report of age at menarche was used to avoid telescoping bias (Pickles et al., 1994), which occurs when individuals remember events as closer to the date of the interview than they actually are. The Pearson correlation between age at menarche reported at Waves I and II ($r = .76$, $p < .001$) was higher than that between Waves II and III ($r = .53$, $p < .001$) and between Waves I and III ($r = .53$, $p < .001$). This may be attributable to telescoping or the fact that menarche at Wave III was reported in years, whereas menarche at Waves I and II was reported in months and years. Ten girls (2% of the sample) were missing data on age at menarche;

they were retained in the sample because they provided data on peer and individual risk-taking.

Self-rated body changes. At Wave I, adolescents' ratings of body changes were assessed using two Likert scale items that asked about breast development (1 = *My breasts are about the same size as when I was in grade school* to 5 = *My breasts are a whole lot bigger than when I was in grade school; they are as developed as a grown woman's breasts*) and body curviness (1 = My *body is about as curvy as when I was in grade school* to 5 = *My body is a whole lot more curvy then when I was in grade school*). Both these items were correlated with age (breast development: $r = .18$, $p < .001$; curviness: $r = .21$, $p < .001$). To account for age differences among participants, scores were standardized within year of chronological age ($M = 0$, $SD = 1$) and the standardized scores for the two items were summed. Thus, higher scores reflect that *an adolescent perceived herself to be more physically developed than her same-age peers perceived themselves*. This is a commonly used technique for transforming a measure of pubertal status to one of pubertal timing in an age-heterogeneous sample (Carter, Silverman, & Jaccard, 2013; Ge, Natsuaki, Neiderhiser, & Reiss, 2007).

Peer comparison. At the Wave I interview, adolescents were asked "how advanced is your physical development compared to girls your age?" Responses fell on a five point scale (1 = *"I look younger than most"* to 5 = "*I look older than most*"). This item was used as a measure of pubertal timing based on peer comparison, with higher scores indicating that *an adolescent perceived herself to be more physically developed than she perceived her peers to be*. This item was correlated with age-standardized self-rated body changes ($r = .36$, $p < .01$). Unlike the items used to assess body changes, this measure was not correlated with age ($r = -.09$, $p = .07$). Moreover, there was

homogeneity of variance across ages ($F(6,476) = .40$, $p = .83$), indicating equivalent variability in the peer comparison measure in older versus younger girls.

Analytic Plan

Initial exploratory analyses were conducted in SAS to assess within- and between-pair correlations between pubertal timing, risk-taking, and peer risk-taking. Subsequent structural equation modeling (SEM) analyses were performed in M*Plus* (Muthén & Muthén. 1998-2010), using Full Information Maximum Likelihood (FIML) to account for missing data. We conducted separate analyses for each measure of pubertal timing (age at menarche, body changes, and peer comparison). Absolute model fit was assessed using root mean square error of approximation (RMSEA) and chi-square values, and nested models were compared using differences in log-likelihood. Age was included as a covariate in predicting individual and peer risk-taking.

First, to assess whether peer risk-taking predicted individual risk-taking, controlling for genetic and environmental selection effects, we fit three multivariate twin models, each of which included peer risk-taking, target risk-taking, and one measure of pubertal timing. This model is illustrated in Figure 5 (for one twin only).

The twin model (see Neale & Cardon, 1992, for a full explanation of the twin model) decomposes variation in a measured phenotype into three sources—additive genetic (A) factors, shared environmental factors (factors that make siblings similar, C), and nonshared environmental factors (factors that make siblings different, plus measurement error, E). Based on genetic theory, the correlation between additive genetic factors is fixed at 1.0 for monozygotic (MZ) twins and 0.5 for dizygotic (DZ) twins. The models used in the current analysis are based on assumptions that may be overly simplistic (Charney, 2012), as recent studies suggest that monozygotic twins are not

genetically identical, due to mutations, postnatal transpositions, somatic mosaicism, and epigenetic effects. However, the aim of the current study was not to estimate heritability but rather to control for genetic and environmental confounds by leveraging the fact that both MZ and DZ twins share an array of genetic and environmental characteristics that are not controlled for in standard epidemiological designs.

Because previous studies with these data have shown minimal effects of shared environment on pubertal timing (Ge et al., 2007; Harden & Mendle, 2012), we fit models in which the shared environmental variance in each measure of pubertal timing was set to 0. Dropping this parameter did not compromise fit for the models of peer comparison ($\Delta\chi^2$ = .47, Δdf = 4, p = .98), body changes ($\Delta\chi^2$ = .1.25, Δdf = 3, p = .74), or age at menarche ($\Delta\chi^2$ = .26, Δdf = 4, p = .97).

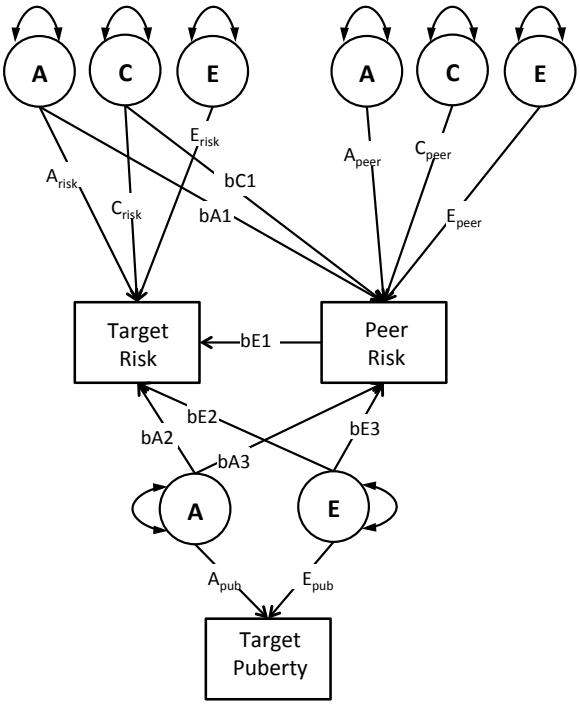


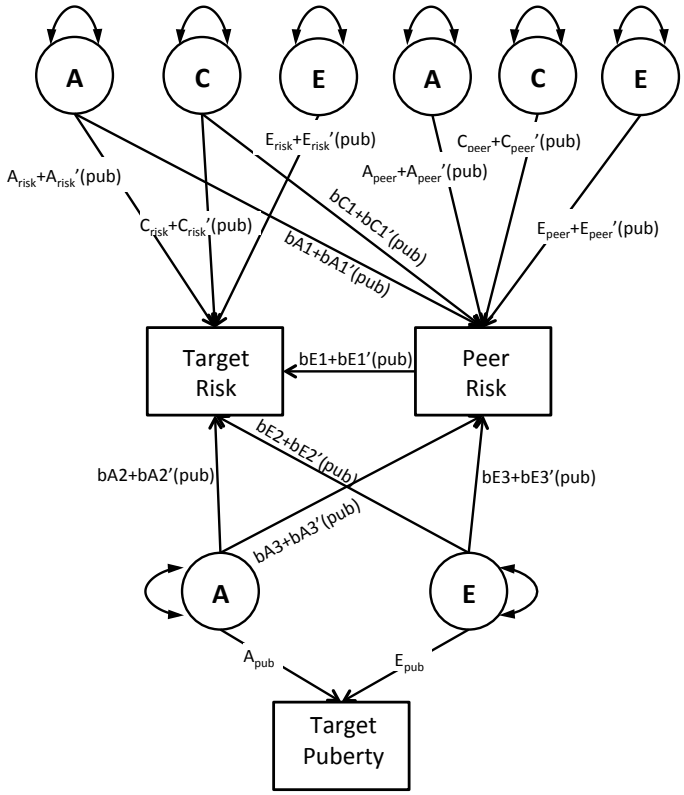
Figure 5. Trivariate Model of Pubertal Timing, Risk-Taking, and Peer Risk-Taking.

Note: A = additive genetic; C = shared environmental; E = non-shared environmental. Only one twin per pair shown. Each measure of pubertal timing modeled separately.

The multivariate model also decomposes associations between phenotypes into genetic, shared environmental, and nonshared environmental pathways. In Figure 5, the path labeled bA1 represents the association between peer and individual risk-taking

that is due to overlapping genetic factors—the extent to which genes that influence propensity for risk-taking also influence one's tendency to associate with risk-taking peers. The path labeled bC1 represents the association between individual and peer risk-taking due to environmental factors that make siblings similar. Environmental factors such as parents, schools, neighborhoods, cultural values, and socioeconomic status may make siblings similar in both their own risk-taking and their peers' risk-taking. Together, paths bA1 and bC1 measure genetic and environmental "selection effects" —the extent to which genetic and environmental factors shared by twins raised in the same home predict both individual risk-taking and affiliation with risk-taking peers. The path labeled bE1 is the association between peer and individual risk-taking that remains after controlling for these selection effects. This path reflects whether the twin who has a riskier peer group engages in more risk-taking than her co-twin. This pathway, referred to as the "quasi-causal" pathway or the "within-pair effect," is how we operationalize peer influence in the current study.

In the second step, to assess whether pubertal timing moderated the association between individual and peer risk-taking, we allowed the cross-paths between these phenotypes to be moderated by each measure of pubertal timing. The moderation model is shown in Figure 6. The terms bA1', bE1', and bC1' represent interactions between pubertal timing and the genetic, nonshared environmental, and shared environmental associations between individual and peer risk-taking. Of central interest in this second step is the term labeled bE1', which is the interaction between pubertal timing and peer influence on individual risk-taking. A significant interaction term would indicate that pubertal timing moderated the influence of risky peer affiliation on individual risk-taking. Analyses were conducted separately for each measure of pubertal timing. For peer comparison and self-rated body changes, higher values reflect earlier pubertal timing; a



positive interaction term bE1' would therefore indicate that girls with earlier pubertal timing are more susceptible to peer influence than later maturing girls. For age at menarche, because higher values reflect later pubertal timing, a *negative* interaction term bE1' would indicate increased susceptibility for early maturing girls.

Figure 6. Interaction Model of Pubertal Timing, Risk-taking, and Peer Risk-taking.

Note: A = additive genetic; C = shared environmental; E = non-shared environmental. Interaction terms denoted by apostrophe. Only one twin per pair shown. Each measure of pubertal timing modeled separately.

RESULTS

Descriptive Statistics

Descriptive statistics and correlations between study variables are shown in Table 6. Because target and peer risk-taking scores were positively skewed, analyses were also performed using log-transformed scores. This transformation did not affect results, and the current study uses the non-transformed scores for both target and peer risk-taking. Peer risk-taking (M = 1.05, SD = .57) was correlated with individual risk-taking (M = .89,

SD = .95), r = .24 (p < .001). The twin pair correlation for risk-taking was r = .40 (p < .01) for DZ twins and r = .51 (p < .01) for monozygotic MZ twins. The twin pair correlations for peer risk-taking were similar for MZ twins (r = .64, p < .01) and for DZ twins (r = .62, p < .01), indicating that shared environment accounted for much of the variance in peer risk-taking; this is likely due to the fact that both MZ and DZ twins attended the same school as their co-twins, and were therefore selecting friends from the same pool. As shown in Table 6, the three measures of pubertal timing—age at menarche, body changes, and peer comparison—were correlated with each other in the expected directions. In addition, all three measures were more highly correlated in MZ twins than in DZ twins, indicating significant genetic variance across all measures of pubertal timing. None of the measures of pubertal development significantly correlated with individual or peer risk-taking.

In addition to risk-taking, other basic characteristics of the peer groups were

Table 6. Correlations and Descriptive Statistics for Key Study Variables

(*N*=248 Twin Pairs)

	Risk Taking	Peer Risk Taking	Age at Menarche	Peer Comparison	Body Changes
Risk Taking		.24**	.05	-.07	.05
Peer Risk Taking			-.06	.02	.09
Age at Menarche				-.33**	-.20**
Peer Comparison					.36**
Twin Pair Correlations	.47**	.62**	.46**	.43**	.30**
MZ Twins	.51**	.62**	.59**	.54**	.35**
DZ Twins	.40**	.64**	.30**	.28**	.24*
Mean (SD)	.89 (.95)	1.05 (.57)	12.31 (1.44)	3.02 (1.01)	0.00 (1.00)

Note: Self-rated body changes were standardized by age. * p < .05, ** p < .01.

examined. The mean proportion of males in the peer group was .41 (*SD* = .22). This was weakly correlated with peer risk-taking ($r = .11, p < .05$) but not significantly correlated with any other measured variables (age, age at menarche, body changes, peer comparison, or individual risk-taking). The average age of the peer group was correlated with age ($r = .88, p < .01$) and with peer risk-taking ($r = .20, p < .01$) but not significantly correlated with any other measured variables.

Does affiliation with risk-taking peers predict individual risk-taking?

Parameter estimates for the multivariate models that were used to address this initial question are shown in Table 7. Separate models were estimated for each of the three measures of pubertal timing: age at menarche, peer comparison, and self-rated body changes.

Age at menarche. The total variance in age at menarche (shown in the first column of Table 2) is estimated by summing the squared paths from the A and E components (labeled A_{pub} and E_{pub}) to age at menarche. The proportion of variance due to genetic and environmental factors can be calculated by dividing each squared path to age at menarche by the total variance in age at menarche. Following this formula, we found variance in age at menarche was due primarily to genetic factors (59%), with remaining variance due to nonshared environmental factors (41%). The regression paths between peer risk-taking and individual risk-taking are labeled bA1, bC1, and bE1. We found that the association between peer and target risk-taking was due to common genetic effects (bA1 = .23, $p < .05$). There were no significant shared or nonshared environmental pathways between peer and individual risk-taking (bC1 = .21, $p = .20$; bE1 = -.29, $p = .28$). Residual variance in peer risk-taking (independent of target risk-taking) was due to shared environmental (52%), and nonshared environmental (48%) factors.

Table 7. Unstandardized parameter estimates from models of pubertal timing, peer risk-taking, and target risk-taking (N = 248 twin pairs)

	Age at Menarche	Peer Comparison	Body Changes
Model Fit Indices			
χ^2 (*df, P*)	72.77	52.94	63.11
	(42, .005)	(43, .14)	(43, .04)
RMSEA	.07	.04	.06
Variance in pubertal timing			
A_{pub}	1.11**	.75**	.64**
E_{pub}	.92**	.69**	.77**
Variance in target risk-taking			
A_{risk}	.61**	.76**	.59**
C_{risk}	.42	.49*	.43
E_{risk}	.69**	.68**	.69**
Variance in peer risk-taking			
A_{peer}	0.00	0.00	0.00
C_{peer}	.34**	.33**	.34**
E_{peer}	.33**	.25**	.33**
Regression Parameters			
Peer risk-taking→Target risk-taking			
Genetic path (bA1)	.23*	.32**	.21**
Shared environmental path (bC1)	.21	.32**	.21**
Nonshared environmental path (bE1)	-.29	-.34	-.30
Pubertal timing→Target risk-taking			
Genetic path (bA2)	0.00	-.07	.15
Nonshared environmental path (bE2)	.03	0.00	-.04
Pubertal timing→Peer risk-taking			
Genetic path (bA3)	-.01	.05	.11
Nonshared environmental path (bE3)	-.04	-.03	-.03

Note: *$p < .05$, **$p < .01$. Variance in target risk-taking indicates variance that is independent from pubertal timing. Variance in peer risk-taking indicates variance that

is independent from pubertal timing and individual risk-taking.

Peer comparison. Variance in the peer comparison measure of pubertal timing was due to genetic factors (54%) and nonshared environmental factors (46%). The first trivariate model found that the genetic, shared environmental, and nonshared environmental paths between peer and individual risk-taking were not significant (bA1 = .20, p = .07; bC1 = .23, p = .16; bE1 = -.25, p = .36). However, none of the paths was reliably different than zero. Given our relatively small sample size, we suspected we did not have sufficient power to differentiate between genetic (bA1) and shared environmental (bC1) selection effects. Consequently, we fit an additional model that set the genetic and shared environmental paths to be equal to each other (following Harden et al., 2008) in order to reduce the number of estimated parameters in our model. We found these paths to be significant and positive (bA1 = bC1 = .32, $p < .01$). Setting these paths to be equal did not compromise model fit ($\Delta\chi^2 = .21$, $\Delta df = 1$, $p = .65$). This suggests that there was a positive association between affiliation with risk-taking peers and individual risk-taking, and that this association was due to unmeasured genetic and/or environmental confounds shared by twins raised in the same family. The nonshared environmental pathway, bE1, was not significant in any of the models (bE1 = -.34, p = .18), indicating that siblings who differed in exposure to risky peer groups did not differ in their own risk-taking.

Self-rated body changes. Variance in self-rated body changes was due to genetic (41%) and nonshared environmental effects (59%). Again, the genetic and shared environmental pathways between peer and individual risk-taking were set to equality and were positive significant (bA1 = bC1 = .21, $p < .01$), whereas the nonshared environmental pathway was not significant (bE1 = -.30, p = .26). Residual variance in

peer risk-taking was again due to shared environmental (52%) and nonshared environmental (48%) effects.

Is the relationship between peer and individual risk-taking moderated by pubertal timing?

Age at menarche. Results of moderation models are shown in Table 8. Again, the primary paths of interest are the regressions between peer and individual risk-taking. Each path had an intercept, and an interaction term that reflected moderation by age at menarche. Age at menarche moderated the nonshared environmental path between peer and individual risk-taking (bE1' = -.48, $p < .01$). The negative interaction term indicated that "peer influence" – the within-twin pair association between peer risk-taking and a girl's own risk-taking – was stronger for girls with earlier age at menarche. In addition, later age at menarche was associated with decreased shared environmental influences (C_{risk}' = .40, $p < .01$) on risk-taking. In sum, results support the hypothesis that girls with earlier pubertal timing, as defined by age at menarche, are more susceptible to peer influence on risk-taking.

Peer comparison. The nonshared environmental pathway between peer and individual risk-taking was moderated by the peer comparison measure of pubertal timing (bE1' = .46, $p < .05$). The within-twin pair association between peer and individual risk-taking was stronger for girls who perceived themselves as more developed than their peers. The genetic and shared environmental pathways, which represented selection effects, were also moderated by peer comparison, in the opposite direction; this was indicated by negative interaction terms on these paths (bA1' = bC1' = -.09, $p < .01$). Thus it appeared that for girls who perceived themselves as more mature than their peers, the association between peer and individual risk-taking was due less to selection and more to socialization. The peer comparison measure of timing also moderated the unique genetic

and environmental variance in risk-taking (i.e., residual variance in risk-taking that was not shared with either pubertal timing or peer risk-taking). Genetic influence on risk-taking was suppressed among girls with earlier pubertal timing (A_{risk}' = -.30, $p < .05$), as was nonshared environmental influence (E_{risk}' = -.12, $p < .05$). Both of these interaction effects were in the same direction, indicating that there is less overall residual variance in risk-taking (independent of puberty and peers) among early-maturing girls.

Self-rated body changes. Unlike age at menarche and peer comparison, self-rated body changes did not appear to moderate the quasi-causal path between peer and individual risk-taking (bE1' = .04, $p = .79$). Thus there was no evidence that associations between peer and individual risk-taking were moderated by age-standardized ratings of body changes.

Table 8. Unstandardized parameter estimates from best-fitting interaction models of pubertal timing, peer risk-taking, and target risk-taking ($N = 248$ twin pairs)

	Age at Menarche	Peer Comparison	Body Changes
Regression parameters			
Peer risk-taking→target risk-taking			
Main genetic path (bA1)	.25**	.49**	.19**
Genetic path x puberty interaction (bA1')	.03	-.09**	.01
Main shared environmental path (bC1)	.33**	.49**	.19**
Shared environment x puberty interaction (bC1')	.03	-.09**	.01
Main nonshared environmental path (bE1)	-.24	-1.61*	-.25
Nonshared environment x puberty interaction (bE1')	-.48**	.46*	.04
Pubertal timing→target risk-taking			
Main genetic path (bA2)	.59*	-.64*	.10
Gene x puberty interaction (bA2')	-.01	.07	-.04
Main nonshared environmental path (bE2)	.52**	-.10	-.06

Table 8, cont.

Nonshared environment x puberty interaction (bE2')	.01	-.09	.04
Pubertal timing→peer risk-taking			
Main genetic path (bA3)	-.01	.11	.11
Gene x puberty interaction (bA3')	0.00	-.02	-.03
Main nonshared environmental path (bE3)	-.05	-.15	-.02
Nonshared environment x puberty interaction (bE3')	.01	.04	.06
Variance in pubertal timing			
Main effect of genes (A_{pub})	1.11**	.75*	.64
Main effect of nonshared environment (E_{pub})	.92**	.69*	.77
Variance in target risk-taking			
Main effect of genes (A_{risk})	-.66**	1.07	.50
Gene x puberty interaction (A_{risk}')	.21	-.30*	.01
Main effect of shared environment (C_{risk})	-.06	1.45**	.52*
Shared environment x puberty interaction (C_{risk}')	.40**	-.26	.32
Main effect of nonshared environment (E_{risk})	.62*	1.03**	.64**
Nonshared environment x puberty interaction (E_{risk}')	-.02	-.12*	-.14**
Variance in peer risk-taking			
Main effect of genes (A_{peer})	0.00	0.00	.23
Gene x puberty interaction (A_{peer}')	0.00	0.00	-.02
Main effect of shared environment (C_{peer})	.17	0.00	0.00
Shared environment x puberty interaction (C_{peer}')	-.029	0.00	0.00
Main effect of nonshared environment (E_{peer})	.34**	.16*	.32**
Nonshared environment x puberty interaction (E_{peer}')	-.04	.05*	.09**

Note: *p < .05, **p < .01

Sensitivity analyses

We conducted three sets of sensitivity analyses to determine whether the broad age range of our sample impacted our main findings for moderation. First, to address the possibility that the peer comparison measure confounded pubertal status and timing, particularly for younger girls, we standardized this measure by age in years (as we had done with self-rated body changes). Findings did not change using this age-standardized measure (bE1' = .72, $p < .001$). Next, following previous studies using Add Health puberty data (Ge et al., 2007; Halpern, King, Oslak, & Udry, 2005), we limited our sample to adolescents ages 12-17 (174 twin pairs) and reran analyses of peer comparison and self-rated body changes. For peer comparison, the moderating effect was of similar magnitude, but the standard errors were larger and the coefficients were at the threshold of conventional statistical significance (bE1' = .55, $p = .05$). Self-rated body changes did not moderate peer influence in this younger sample (bE1' = -.08, $p = .72$). Next we restricted the age range to adolescents ages 14-19 (212 twin pairs). Again we found a significant moderating effect for peer comparison (bE1' = .93, $p < .001$) and not for self-rated body changes (bE1' = .18, $p = .34$). Finally, we reran moderation analysis for age at menarche, limiting the sample to girls who had started menstruating by Wave I, and found a moderating effect in this restricted sample (bE1' = -.53, $p < .01$).

In a final set of analyses, we examined whether the moderation effect was driven by racial or ethnic differences in pubertal timing. We ran two additional models, one including dummy-coded race (African American vs. non-African American) as a moderator and one including ethnicity (Hispanic vs. non-Hispanic) as a moderator. There were no moderating effects of race (bE1' = -.14, $p = .79$) or ethnicity (bE1' = -.35, $p = .79$) on the quasi-causal association between peer and individual risk-taking. Thus, the

moderating effects of pubertal timing could not be explained by racial or ethnic differences in pubertal timing.

DISCUSSION

The current study used a quantitative behavior genetic design to examine whether girls' pubertal timing moderated the quasi-causal association between peer risk-taking and individual risk-taking. Results suggest girls with earlier ages at menarche and girls who perceived themselves to be more physically developed than their peers were more susceptible to peer influence on risk-taking. These findings align with previous studies showing that pubertal timing moderates the adverse effects of deviant peer affiliation (e.g. Fergusson et al., 2007; Costello, 2007). This pattern is consistent with the "contextual amplification hypothesis" (Ge et al., 2002; Stattin et al., 2011), which proposes that earlier pubertal maturation may magnify the effects of contextual factors such as peers, neighborhoods, schools, and families. These findings are also congruent with recent behavior genetic research that has found that more of the variation in delinquency can be attributed to environmental differences among girls with early pubertal timing (Burt et al., 2006; Harden & Mendle, 2012). To our knowledge, no other studies have used quantitative genetic methods to examine pubertal timing as a moderator of the relation between peer-reported and self-reported delinquency.

The age range of this sample (mean age = 16 years) has important implications for interpreting results. A sample of middle adolescents is ideal for studying risk-taking, which increases markedly during this period and is a leading cause of morbidity and mortality for this age group (CDC, 2009). Yet by middle adolescence most girls are "post-pubertal" (or at least post-menarcheal). Thus, we are not necessarily examining the effects of earlier pubertal timing during the period when most of the physical changes of

puberty are occurring. Rather, these effects represent the persisting effects of individual differences in pubertal timing into later adolescence. We are not the first to find that measures of pubertal timing have predictive power beyond the age at which girls are initiating puberty (Graber et al., 1997; Haynie, 2003; Ge et al., 2007). Indeed, some have suggested that studies of older adolescents are necessary to clarify the effects of pubertal timing; as Angold and Costello (2006) noted, "it is not until after puberty, when everyone is fully mature, that the timing of any pubertal event is unambiguously unconfounded" with pubertal status (p. 924). These associations nevertheless pose a developmental enigma: Why does early age at menarche still matter if everyone is post-menarcheal? What does it mean for a girl to "look older" (or think she looks older) than her peers if she and they would likely all be considered reproductively mature on a Tanner-stage scale? Our conjecture is that pubertal events occurring in late childhood and early adolescence have relatively enduring effects on girls' self-perceptions, and that these differences in self-perception represent an important mechanism for how effects of pubertal timing persist into later adolescence.

There were two subjective measures used in the current study: perceived body changes standardized by age, and a peer comparison measure. Perceived development compared to one's peers moderated the association between peer and individual risk-taking, whereas perceived body changes did not. There are several possible explanations for these discordant findings. These subjective measures (body changes and peer comparisons) reflect different aspects of pubertal timing and probably other constructs as well. The peer comparison measure ("how advanced is your physical development compared to other girls your age") likely reflects characteristics of an adolescent's reference group—the group of "other girls your age" that forms the basis for self-comparison. In addition, the response options to this item are somewhat ambiguous,

ranging from "I look younger than most" to "I look about average" to "I look older than most." Many other factors besides physical development contribute to one's apparent age, including attempts to manipulate apparent age with clothing, makeup, or tattoos. Thus, this measure of peer-normative pubertal timing, while correlated with age at menarche and breast development, also taps constructs such as perceptions of peers and impression management, which may be closely linked with susceptibility to peer influence.

The lack of moderation by self-rated body changes suggests that a girl's perception of how her body has changed may be less influential—at least on susceptibility to peer influence on risk-taking—than her perception of how her body differs from her peers. This pattern of results– in which significant effects were observed for age at menarche and peer comparison but not body changes– was also observed in a study of dieting in adolescence (Harden, Kretsch, & Mendle, 2012). It is also possible that ratings of body and breast development are confounded by body type and weight. This may be particularly true of the Add Health participants, many of whom were in mid-adolescence and post-menarcheal at the start of the study. Therefore, a response of *"my body is a lot more curvy than when I was in grade school"* may be more indicative of one's body type than of one's pubertal stage.

A commonly proposed explanation for the link between pubertal timing and risk-taking—whether it is conceptualized in terms of mediation or moderation—is that girls who experience early pubertal timing and who appear older than their same-age peers affiliate with older, more male-dominated peer groups. Because they are younger in chronological age, they may rely more on their older peers for opportunities to engage in risky behavior. However, current findings are unlikely driven solely by the age and gender of one's friends, because phenotypic correlations between pubertal timing and

peers' age and gender showed no significant associations. In addition, the moderating effects of pubertal timing found in the current study were not driven by racial or ethnic differences in pubertal timing. However, given the comparatively small number of racial and ethnic minority twin pairs in Add Health, and the disproportionate number of Hispanic girls with no identifiable peer nominations, we did not have sufficient power to test a three way interaction (peers × race and ethnicity × pubertal timing). The fact that Hispanic girls had fewer identifiable peer nominations suggests there may be qualitative and quantitative ethnic differences in peer group composition. There may be important racial and ethnic differences in the relation between pubertal timing and peer influence that are not captured in our study. Cavanagh (2010) found an interaction between peer academic achievement and pubertal timing that predicted age at first sex in Hispanic girls only, and found no significant associations between friendship group characteristics, pubertal timing, and age at first sex for African American girls. Understanding the extent to which race and ethnicity may impact susceptibility to peer influence via differences in pubertal timing remains an important question of future study.

Our findings should be interpreted in light of several limitations. First, the sample size is small compared to a typical twin sample, which is important given the complexity of the models tested in this study. Power was also limited due to the substantial portion of missing data on peer risk-taking. Second, other characteristics of the study sample limit the extent to which we can generalize these findings. It is possible that different associations would be found in a higher risk sample or in a sample of males, who engage in more risk-taking on average. Research on pubertal timing in males is limited, likely due to the lack of a "male equivalent" to females' age at menarche, but recent studies suggest pubertal timing has important effects on psychosocial outcomes in males (reviewed in Mendle & Ferrero, 2010), and that peer relationships are key mechanisms

for the psychosocial impact of pubertal development in males (Mendle, Harden, Brooks-Gunn, & Graber, 2012). In particular, a recent longitudinal study using data from Add Health examined pubertal timing (measured by age-standardized ratings of body changes), alcohol use, and peer alcohol use, and found an interaction between pubertal timing and peer alcohol use that was *only* present among males (Biehl et al., 2007). Further exploration of these gender differences, using longitudinal quasi-experimental designs and peer-reported risk-taking, is warranted. Indeed, given current evidence that the subjective peer comparison measure taps a meaningful psychological construct, the absence of a more objective pubertal milestone should not hold us back from exploring similar questions about puberty, peers, and risk-taking in males.

Third, because the current study used cross-sectional data, the direction of peer influence is not clear. We arbitrarily designated the twin girls in this sample as targets of influence, but it is also likely that they are sources of influence. Thus an additional explanation of our findings is that girls with earlier pubertal timing exert more influence on their peers. These dual interpretations are not necessarily mutually exclusive. Research on peer influence susceptibility relies on a distinction between "leaders" and "followers," (Allen, Porter, & McFarland, 2006), but it is plausible that individuals oscillate between these roles as they navigate the social transition of adolescence.

Finally, the current study used social network data but did not involve formal social network analysis or statistically control for peer groups nested within schools. Social network analytic methods offer a chance to model selection and influence simultaneously (Snijders, 2010), and, when applied to genetically-informative longitudinal data, these tools can answer important questions about susceptibility to peer influence. The fact that all twins (both MZ and DZ) were selecting peers from the same pool inflates shared environmental selection effects over genetic selection effects. It also

provides a rigorous quasi-experimental test of peer influence: twins are matched on age and grade, and have equal opportunity to select friends from the same pool of grademates.

The preponderance of research on peer influence suggests that both selection and socialization shape adolescent behavior. Using a behavioral genetic model of peer influence, this study suggests both processes are moderated by pubertal timing. In recent years, early pubertal timing has been the focus of both scientific study and popular media (e.g., Weil, 2012), particularly given observed trends in earlier pubertal development. These findings highlight the importance of considering the peer context in which biological changes occur and of using multiple measures of pubertal timing. Key areas for research include exploration of gender, race, and ethnicity as additional moderators of peer influence and incorporating longitudinal social network methods.

Chapter 4: Peer group similarity in perceptions of pubertal timing[2]

BACKGROUND

The timing of puberty is a consistent predictor of a number of health behaviors and associated outcomes. Among girls, early pubertal timing is associated with increased risk for depression, disordered eating, risky sexual activity, sexual victimization, early childbearing, delinquency, and substance abuse (reviewed in Graber, Seeley, Brooks-Gunn, & Lewinsohn, 2004). Less research has examined the role of pubertal timing in boys' adjustment and findings are more varied. The association between boys' earlier pubertal timing and externalizing pathology appears robust (reviewed in Mendle & Ferrero, 2012). Regarding internalizing problems, however, some longitudinal studies have found higher rates of depression among early maturing boys (e.g., Rudolph & Troop-Gordon, 2010), whereas other studies suggest higher rates of depression in later maturing boys, particularly in the context of problematic peer relationships (Conley & Rudolph, 2009).

Much of the research on health sequelae of pubertal timing has relied on self-report measures of pubertal timing, which correspond only modestly with clinical measures of development such as physical exam (Dorn & Biro, 2011). Self-report measures of physical development are considered imperfect proxies of actual development, and their use is often noted as a limitation in studies of puberty (e.g., Shirtcliff, Dahl, & Pollak, 2009; Dorn et al., 2006). Age at menarche is considered the most "objective" self-report measure of timing in girls, because it is a discrete event that tends to be reported with reasonable accuracy compared to historical medical records (Casey et al., 1991). However, recall of age at menarche is not perfect, and menarcheal

---

Kretsch, N., Mendle, J., Cance, J. D., & Harden, K. P. (in press). Peer Group Similarity in Perceptions of Pubertal Timing. *Journal of Youth and Adolescence*. The first author conceived of the study, performed statistical analysis, and drafted the manuscript.

status is a dichotomous measure that captures only one aspect of pubertal development occurring relatively late in the process (Shirtcliff et al., 2009). Nevertheless, these subjective measures of pubertal timing are often stronger predictors of key health outcomes, such as eating disorders, substance abuse, and delinquency than more "objective" indicators (e.g., Deppen et al., 2012). Given the impact of subjective pubertal timing on adolescent behavior, it is important to understand the mechanisms that may lead an adolescent to perceive her or himself as on-time or off-time. This article considers the social comparisons and relationships that may contribute to adolescents' self-perceptions of physical maturation.

Measures of Subjective Pubertal Timing

Although often considered together, there are important conceptual distinctions among different types of self-report measures of pubertal timing (Cance et al., 2012). The widely used Pubertal Development Scale (PDS; Peterson, Crockett, Richards, & Boxer, 1988), for example, asks adolescents to rate the progression of specific physical changes (e.g., breast development, voice changes, body hair) using scales with visual or descriptive anchors. Self-report measures can be standardized by age within each sex to create an indicator of pubertal timing (e.g., Natsuaki, Biehl, & Ge, 2009). In the current article, we refer to these age-standardized ratings of body changes as "*stage-normative timing*." Stage-normative timing predicts substance use (Costello, Sung, Worthman, & Angold, 2007), disordered eating (Baker, Thornton, Lichtenstein, & Bulik, 2012), delinquency (Harden & Mendle, 2012), and depression (Natsuaki, Biehl, & Ge, 2009). Stage-normative measures are modestly correlated with physician ratings of Tanner stages and with pubertal hormones (Bonat et al., 2002).

In contrast, "peer-normative" or "relative" pubertal timing measures directly ask adolescents to compare themselves to their peers. Some of the strongest effect sizes for the most clinically significant outcomes are found in studies that explicitly ask one to compare their physical development to that of their peers. For example, a study that linked pubertal timing with risk for suicide assessed timing by asking "when you look at yourself now, do you think that you are more or less physically mature compared to others [of the same sex] of your age?" (Wichstrom, 2000). In the current article, we refer to this type of peer comparison measure as "*relative pubertal timing*." Earlier relative pubertal timing is associated with internalizing and externalizing problems in both males and females (Carter, Caldwell, Matusko, Antonucci, & Jackson, 2011; Yuan, 2007). These associations persist into mid and late adolescence, even after pubertal development is complete (Kretsch, Mendle, & Harden, 2014).

Social Comparisons and Subjective Pubertal Timing

Identity development is a key task of adolescence, and the social context—specifically the peer context—plays a role in this process (Erikson, 1950; McAdams & Olson, 2010). Individuals compare themselves to significant others in order to make sense of who they are and where they fit in the world (Finkenauer et al., 2002; Adams & Marshall, 1996). This process is made explicit in survey measures of relative pubertal timing, which introduce "a psychosocial component that is missing from the stage-normative measure" (Cance et al., 2012, p. 766). To our knowledge, however, no previous studies have examined whether the peer context is differentially related to stage-normative versus relative pubertal timing. If relative pubertal timing measure reflects a peer comparison process that stage-normative measures does not, the pubertal timing of

one's peer group is hypothesized to predict relative pubertal timing more strongly than stage-normative timing.

Peers' pubertal timing may relate to individuals' perceived relative pubertal timing in a number of ways. Social comparison theory describes two processes that may occur when individuals compare themselves to others: contrast and assimilation (Festinger, 1954; Blanton, 2001). When adolescents are asked to compare their development "to other girls or boys your age," their reference group may be limited to the peers who are visible to them, particularly, those peers with whom they frequently interact. If a peer contrast process is unfolding, peers' pubertal timing will influence an individual's ratings such that an early-maturing adolescent may not perceive himself or herself as early maturing if his/her friends are similarly developed, and those who mature on-time or late according to population norms might perceive themselves as early developers if they matured earli*er* than their peer group. Peer contrast would thus result in an inverse association between peer pubertal timing and one's own perceived pubertal timing for early maturing adolescents.

Alternatively, it is possible that friends and schoolmates will be similar in perceptions of pubertal timing, through processes of selection or assimilation. Early or late maturing adolescents may experience rejection by or conflict with their typically-developing peers (e.g., Haynie & Piquero, 2006), and, as a result, select friends who are similar in pubertal status. Through assimilation, adolescents also incorporate features of their peer group into their self-image. Thus, adolescents who have early maturing friends may perceive themselves (or desire to see themselves) as early maturing, regardless of their objective pubertal status. This may play a role in similarity for behaviors such as substance use, delinquency, athletic involvement, academic achievement, and sexual

activity, all of which are associated with pubertal timing (reviewed in Mendle, Turkheimer, & Emery, 2007; Mendle & Ferrero, 2012).

Cross-cultural differences may also play a role in perceived development. Friends and schoolmates are often similar to one another in race and ethnicity (McPherson, Smith-Lovin, & Cook, 2001), and there are racial and ethnic differences in pubertal timing: on average, non-Hispanic White adolescents mature slightly later than Black and Hispanic adolescents (Sun et al., 2002), although they do not necessarily perceive themselves this way (Cance et al., 2012). Friends are also similar in broader aspects of physical appearance, including body mass index (Cohen-Cole & Fletcher, 2008; Fowler & Christakis, 2008), which is linked with both perceived and actual earlier development and varies across ethnic and socioeconomic groups (Bonat et al., 2002).

Moderation by Individual Characteristics

Studies examining the health effects of pubertal status often limit samples to a narrow age range in early adolescence (i.e., middle school), when there is the greatest variation in objective pubertal status (Lee & Styne, 2013; Parent et al., 2003). By mid-adolescence (i.e., high school), when nearly everyone has reached Tanner Stage 5,one might expect less variation in perceived relative timing, with most adolescents rating themselves as average. Previous analyses, however, have shown that there are individual differences in perceived relative timing even in high school age samples, and that these individual differences continue to predict important behavioral outcomes (e.g., Kretsch et al., 2014). It is not yet clear whether the social influences for perceived relative timing are the same for younger adolescents, for whom the physical changes of puberty are still ongoing, as they are for older adolescents, whose objective physical development is essentially complete. On the other hand, several lines of research—on puberty, peer

influence, and social cognition—suggest that peer pubertal timing might be more salient for younger adolescents. Perceptions of pubertal timing become more stable (Cance et al., 2012) and more accurate (Dubas, Graber, & Petersen, 1991) over time, suggesting that, by late adolescence, ratings of relative timing will be less influenced by an individual's peer group. Some research suggests that susceptibility to peer influence decreases between ages 14 and 18 (Steinberg & Monahan, 2007). If peer influence on perceptions of relative pubertal timing show the same pattern, one would expect that associations between self-reports of individual and peer timing would decrease over time.

In addition to age, it is important to consider potential gender differences in how peers shape perceptions of pubertal timing. Gender differences in social orientation, peer relationships, peer network structure, and body comparison tendencies suggest that peer comparisons may differ for males and females. Studies of friendship network structure suggest that girls have more intimate friendship networks and are more connected to school peer networks than boys (Urberg, Degirmencioglu, Tolson, & Halliday-Scher, 1995). A study on social comparison and body image in 7th and 10th graders found that girls made more appearance-based social comparisons than boys (Jones, 2001). Similar gender differences have been observed among first-year college students (O'Brien et al., 2009). Given these findings, and the longstanding view of girls as more peer-oriented than boys (e.g., Rose & Rudolph, 2006), one might expect that peer characteristics would be more salient for girls' self image that others. However, research on gender differences in susceptibility to peer influence in general is inconclusive. As Brechwald and Prinstein (2010) summarized, gender moderates peer socialization effects "only within more complex two- and three-way interaction terms that also consider age and the specific behavior being influenced" (p. 172).

Goals and Hypotheses of the Current Study

The goals of the current study were twofold. First, in an exploratory analysis, we examined whether adolescents were similar to their friends and schoolmates in three measures of pubertal timing: (a) "stage-normative" pubertal timing, an age-standardized rating of specific body changes; (b) relative pubertal timing, which directly asked adolescents how developed they were compared to their same-age peers; and (c) girls' age at menarche. Second, we tested peers' pubertal timing (relative, stage-normative, and age at menarche) as a predictor of individuals' relative and stage-normative pubertal timing. We predicted that, controlling for one's own stage-normative timing (and, for girls, age at menarche), the stage-normative timing of one's peers would predict one's own relative pubertal timing. Specifically, adolescents whose peers reported earlier stage-normative timing would report later relative pubertal timing. This prediction was based on the theory that the relative timing measure elicits a social comparison process in which adolescents use their peers as reference groups. We also predicted that, **c**ontrolling for one's own relative timing, the relative timing of one's peers would *not* predict one's own stage-normative pubertal timing. This prediction was based on the theory that the stage-normative measure does not elicit the same peer comparison process and, as such, should not be influenced by pubertal timing of one's peers. We examined both the main effects of age on perceptions of pubertal timing and the interactions between age and peers' stage-normative pubertal timing. We predicted that peers' stage-normative timing would be especially relevant for younger adolescents' relative pubertal timing, because younger adolescents are in the midst of pubertal changes and may be more attuned to differences in maturation among their peer group. Finally, we performed separate analyses for boys and girls, to explore how the associations between peer and individual perceived pubertal timing differed between genders. We did not test gender as a moderator in the full

sample, given the different measures of pubertal timing for boys and girls (necessitating that all boys would be missing-by-design on the age at menarche variable).

METHOD

Participants

Data were drawn from the National Longitudinal Study of Adolescent Health (Add Health). Add Health includes four waves of data on health and risk behavior in a nationally representative sample of adolescents who were in grades 7-12 at the initial wave in 1994. Add Health used a school-based sampling procedure that started with identifying all schools in the US that had at least 30 students (N = 26,666). Schools were stratified according to region, urbanicity, racial composition, school size, and school type. A random sample of 80 schools was selected from these strata, and invited to participate. The feeder middle schools for the high schools in this sample were also invited. Of the selected schools, 79% agreed to participate, yielding a sample of 134 schools.

A confidential survey was administered to all students in participating schools (N = 90,118) during the 1994-95 academic year. The survey included questions about demographics, academic achievement, school activities, and delinquent behavior. A subsample of 20,745 students was selected to complete a longer, in-home interview between April and December 1995 (Wave I). The home interview included more questions about sensitive topics including sexual activity, drug and alcohol use, and pubertal development. Individuals who completed this interview were interviewed again approximately one year later (Wave II). Two additional interviews, Wave III in 2001-02, and Wave IV in 2007-09, have been completed. A complete description of the Add Health study is available at http://www.cpc.unc.edu/projects/addhealth. The current study uses data primarily from the Wave I in-home interview.

One of the goals of Add Health was to understand adolescent behavior in the context of social networks. Adolescents were asked to nominate up to five male and five female friends and were asked to list best friends first. They were allowed to nominate romantic partners and asked to indicate which, if any, nominated friends were romantic partners. Based on these nominations, it was possible to link data between individual participants and their nominated friends within the same school. There were 16 schools in which all adolescents in the school (rather than a subsample) were recruited for the in-home interview. This sample (N = 3,702) is considered the *saturation sample*. Of the students in the saturation sample, 78% (N = 2,817) had data on friends' pubertal timing because they nominated at least one identifiable, same sex friend from the school roster, and this friend reported his or her own pubertal timing. The current study includes data from these 2,817 individuals (M_{age} = 16.60, range 12.5—20.7) and their same-sex friends.

There were several reasons why some adolescents had no data on friends' pubertal timing (i.e., reasons they were removed from the saturation sample of 3,702). First, they may not have nominated any same-sex friends (N = 232). Second, they may have nominated only out-of-school same-sex friends or friends who were not on the school roster (N = 484). Third, they may have nominated an identifiable friend, but this friend did not provide data on pubertal timing (N = 169). Adolescents without data on friends' pubertal timing were more likely to be Black (χ^2 = 38.18, df = 1, p < .001; 26% vs 16%) and/or Hispanic (χ^2 = 9.46, df = 1, p < .001; 21% vs. 16%) than those with identifiable nominations. There were no differences in age or gender between those with or without data on peer pubertal timing. Demographic differences between the full Add Health sample (N = 20,745) and the current study sample (N = 2,817) are shown in Table 9.

Table 9. Descriptive Statistics for Study Sample (N = 2,817) and Full Add Health Sample (N = 20,745)

	Study Sample Mean (*SD*)	Full Add Health Sample Mean (*SD*)
Age*	16.61 (1.56)	16.16 (1.72)
Female	48.92%	50.52%
Hispanic*	20.37%	17.04%
Black*	15.91%	23.22%
Relative Pubertal Timing	3.18 (1.10)	3.19 (1.13)
Stage-Normative Pubertal Timing	-0.02 (.98)	0.00 (1.00)
Age at Menarche	12.22 (1.40)	12.17 (1.42)
Friends' Mean Age	16.60 (1.42)	
% Hispanic friends	16.89 (32.78)	
% Black friends	11.21 (27.91)	
Friends' Relative Pubertal Timing	3.19 (.88)	
Friends' Stage-Normative Pubertal Timing	.02 (.78)	
Friends' Age at Menarche	12.26 (1.10)	
School Mean Age	16.62 (.96)	
School % Hispanic	14.84 (17.41)	
School % Black	24.12 (19.40)	
School Relative Timing	3.18 (.21)	
School Stage-Normative	-.02 (.25)	
School Age at Menarche	12.22 (.19)	

Note: School- and friend-level variables are only available for the saturation sample. *Study sample differs from full sample at $p < .05$.

Measures

Individual characteristics

Relative pubertal timing. At Wave I, participants were asked, "how advanced is your physical development compared to boys/girls your age?" Response options fell on a five point scale, including: "I look younger than most" (1); "I look younger than some (2); "I look about average" (3); "I look older than some" (4); "I look older than most" (5). The mean response was 3.18 (*SD* = 1.10).

Stage-normative pubertal timing. At Wave I, females were asked to rate breast development (1 = "My breasts are about the same size as when I was in grade school" to 5 = "My breasts are a whole lot bigger than when I was in grade school; they are as developed as a grown woman's breasts") and body curviness (1 = "My body is about as curvy as when I was in grade school" to 5 = "My body is a whole lot more curvy then when I was in grade school"). Males were asked to rate underarm hair growth (1 = "I have no hair at all" to 5 = "I have a whole lot of hair that is very thick, as much hair as a grown man"), facial hair growth (1 = "I have a few scattered hairs, but the growth is not thick," to 5 = "The hair is very thick, like a grown man's facial hair") and voice changes (1 = "it is about the same as when you were in grade school" to 5 = "it is a whole lot lower than when you were in grade school; it is as low as an adult man's voice"). For each item, we calculated each participant's standardized ($M = 0$, $SD = 1$) deviation from the mean response to this item by adolescents of the same age and gender. These standardized scores on each item were summed. Thus, positive higher values reflected more advanced physical development compared to the physical development reported by same-age same-sex adolescents.

Age at menarche. At Waves I and II, female participants were asked "Have you ever had a menstrual period?" If affirmative, they were asked the month and date of their first menstrual period. At Wave III, participants were asked if they had ever had a menstrual period and how old they were when they had their first menstrual period. The current study used the first reported age at menarche for each adolescent (i.e., if an adolescent initially reported that she began menstruating at Wave I and also reported age at menarche at Wave II, the Wave I initial report was used). This was to avoid telescoping bias (Janssen, Chessa, & Murre, 2006), which occurs when individuals remember events as closer to the date of the interview than they actually occurred. The

mean age of menarche in the sample was 12.23 years (SD = 1.40, range = 7 years – 18 years). The Pearson correlation between age at menarche reported at Waves I and II (r = .76, p < .001) was higher than that between Waves II and III (r = .53, p < .001) and between Waves I and III (r = .53, p < .001). This may be attributable to telescoping or the fact that menarche at Wave III was reported in years, whereas menarche at Waves I and II was reported in months and years.

Friend characteristics

Using peer nominations, we identified a group of male friends and female friends for each adolescent. We averaged the pubertal timing measures of each adolescent's same-sex friends to obtain three measures of friends' pubertal timing: friends' relative pubertal timing, friends' stage-normative timing, and female friends' average age at menarche. We also calculated the average age of nominated friends and the proportions of friends who were Black and Hispanic.

School characteristics

We computed the average relative and stage-normative timing reported by males and females at each school, as well as the average age at menarche of females at each school. Each school had an administrator fill out a questionnaire about school characteristics, including the percentages of Black, White, and Hispanic students at each school. The percentages of Black students ranged from 0 % to 99%. The percentages of Hispanic students ranged from 0% to 43%. We also calculated the average age of students at each school, which ranged from 13.92 to 17.29 (M = 16.62, SD = .97).

Analytic Plan

Analyses were performed in SAS v. 9.2. There was no available measure of male pubertal timing that was equivalent to girls' age at menarche; therefore, separate analyses

were performed for males and females. Continuous variables were mean-centered. In an initial step, we examined friend- and school-level similarities for each measure of pubertal timing, by calculating partial correlation coefficients for adolescents and their nominated friends and school-level intraclass correlations.

Relative pubertal timing was then analyzed as the focal dependent variable in a series of regression models that first controlled for individual characteristics and then regressed relative pubertal timing on all measures of peer pubertal timing.

Model 1 included individual characteristics only: self-reported stage-normative timing, age, age-squared, age at menarche (for females), race, and ethnicity. Individuals' stage-normative timing and age at menarche were included to examine whether, after controlling for the expected concordance between these self-report measures, the stage-normative pubertal timing of one's peers would influence a participant's perceived relative pubertal timing. Age and age-squared were included because we used an age-heterogenoues sample and the effect of age on the key measures of pubertal development may not be linear (i.e., the relation between age and puberty may be stronger at younger ages). Model 2 added characteristics of the adolescent's nominated friends as predictors of individuals' relative pubertal timing. Of key interest were the associations between individuals' relative timing and their peers' stage-normative timing. Model 3 added school-level characteristics: the mean level of stage-normative and relative timing reported by same-sex students at each school, as well as demographic characteristics of the schools (percent Black, percent Hispanic, and average age). Subsequent models tested for interactions between age and individual, friend, and school characteristics. Model 4 added an age-by-stage-normative timing interaction and, for females, an age-by-age at menarche interaction. Models 5 and 6 added age-by-friends' pubertal timing and age-by-school timing interactions.

Next, stage-normative timing was analyzed as the focal dependent variable in a similar series of regression models. We regressed stage-normative pubertal timing on all measures of peer pubertal timing. The same approach was used in adding individual, friend, and school characteristics followed by a series of age interaction terms.

For both sets of pubertal timing outcome measures, analyses were initially performed using mixed-effects models using PROC MIXED in SAS to account for potential school-level clustering in pubertal timing. However, for models predicting relative timing, after adding the school-level characteristics in Model 3 the school-level random effects were reduced to zero, indicating that all of the school-level clustering for relative timing was due to these characteristics, yielding a general linear (OLS regression) model with the same results. For models predicting stage-normative timing, after adding the school-level characteristics in Model 3, the school-level random effects *were* significant, so analyses for stage- normative timing used mixed-effects models.

RESULTS

To what extent are adolescents similar to their friends and schoolmates in self-reported pubertal timing?

Results of correlational analyses are shown in Table 10. For each measure of pubertal timing, we examined the partial correlation between self and friends' self-report, controlling for age, race, and ethnicity. Males were similar to their nominated friends in self-reported stage-normative pubertal timing ($r = .16, p < .001$). Females were similar to their nominated friends in relative pubertal timing ($r = .07, p < .05$).

We also estimated the intraclass correlation coefficient (ICC) within each school, for each measure of pubertal timing. The ICC was computed using linear mixed models that estimated the random effect of school, controlling for fixed effects of school demographics (mean age and percentages of Black and Hispanic students). There were

modest between-school differences in male stage-normative pubertal timing (ICC = .04, $p < .05$). However, there were no between-school differences in the other measures of pubertal timing that were not explained by school average age and by racial and ethnic demographics. That is, any apparent clustering of pubertal timing within schools was due to clustering of racial/ethnic minorities within schools.

Table 10. Correlations between individual and same-sex peers' reports of pubertal timing. $N = 2{,}817$.

	Female Relative Pubertal Timing[c]	Female Stage-Normative Timing[d]	Age at Menarche	Male Relative Pubertal Timing[c]	Male Stage-Normative Timing[e]
Partial correlations between adolescents and same-sex friends[a]	0.07*	0.03	0.07	0.06	0.16***
School-level intraclass correlations[b]	0.00	0.00	0.00	0.00	0.04*

*$p < .05$, **$p < .01$, ***$p < .001$

[a]Partial correlations control for peers' age (mean age of nominated peers), race (% of peers who report Black race), and ethnicity (% of peers who report Hispanic ethnicity).

[b]School ICCs control for fixed effects of school racial composition and mean student age.

[c]Relative pubertal timing was assessed by the question "How developed are you compared to other boys/girls your age?"

[d]Female stage-normative pubertal timing is based on age-standardized ratings of breast growth and body curvature

[e]Male stage-normative pubertal timing is based on age-standardized ratings of voice changes, facial hair, and body hair

Does the stage-normative timing of the peer group predict relative pubertal timing?

Results of regression analyses using relative pubertal timing as the dependent variable are shown in Table 11 for males. Model 1 showed the expected positive association between self-reported relative and stage-normative timing ($\beta = .42, p < .001$). Males with greater age-standardized ratings of body changes (earlier stage-normative timing) also reported that they were more physically developed than their peers. Age was negatively associated with relative timing ($\beta = -.56, p < .05$). There was also a significant, negative quadratic effect of age on relative timing ($\beta = -.02, p < .05$): as boys aged, they perceived their pubertal development to be earlier until age 15 when this perception shifted to reflect on-time development. There were no significant racial/ethnic differences in relative pubertal timing among boys.

Model 2 added main effects of the adolescent's nominated friends. Contrary to predictions, there were no associations between a boy's relative pubertal timing and either measure of pubertal timing reported by his nominated friends, after controlling for his own individual characteristics. Main effects of individual stage-normative timing remained significant, as did the quadratic age effect, but the linear effect of age became non-significant. There was a positive association between friends' mean age and relative pubertal timing ($\beta = .09, p < .05$). Males with older friends perceived themselves as more physically developed than others their age. There were no effects of friends' race or ethnicity.

Model 3 added school-level predictors of relative pubertal timing. At the school level, there was a positive association between relative timing and the average relative timing reported by one's schoolmates ($\beta = .85, p < .05$), suggesting school-level similarity for relative timing. This result contrasted with the minimal school-level ICC estimated in our preliminary correlational analyses; the discrepancy may be due to the

additional individual-level and friend-level covariates that were included in the regression models. The racial and ethnic composition of the school was not associated with boys' relative pubertal timing. Effects of age and friends' mean age remained significant in this model. Subsequent models added interactions between age and individual (Model 4), friend (Model 5) and school (Model 6) characteristics. We found no significant interactions between age and any of these characteristics.

Parallel analyses for females are shown in Table 12. Model 1, which included individual characteristics only, showed positive associations between stage-normative and relative pubertal timing ($\beta = .39$, $p < .001$). Age at menarche was also negatively associated with relative pubertal timing ($\beta = -.15$, $p < .001$). Thus, as would be expected, females with earlier age at menarche and with more advanced age-standardized ratings of body changes also rated themselves as more developed than others their age. There were also racial and ethnic differences in relative pubertal timing. Compared to White females, Black ($\beta = -.27$, $p < .001$) and Hispanic ($\beta = -.03$, $p < .001$) females tended to rate themselves as less developed than their same-aged peers.

Models 2 and 3 added characteristics of female friends and schoolmates, respectively. There was no evidence for an association between relative timing and friends' or schoolmates' pubertal timing, and there were no friend-level or school-level effects of age, race, or ethnicity (that is, there were no significant effects of friend or schoolmate race, ethnicity, or age on females' self-reported relative pubertal timing). The individual-level effect of Hispanic ethnicity became non-significant when friend-level effects were added.

Model 4 added interactions between age and individuals' stage-normative timing and age at menarche in predicting relative pubertal timing. The negative association between age at menarche and relative pubertal timing was weaker for older girls ($\beta = .03$,

$p < .05$). Model 4a added interactions between age and Black race and between age and Hispanic ethnicity to test whether these timing × age interactions remained significant when race/ethnicity × age interactions were added. The age × menarche interaction was no longer significant with these additional interaction terms. There was a significant interaction between age and Black race ($\beta = -.23$, $p < .01$), suggesting that the tendency for Black girls to report later relative pubertal timing increased with age. In Model 5, which added interactions between age and friends' pubertal timing, the age × menarche interaction was significant, as was the age × Black race interaction. Overall, these interaction models suggested that perceptions of relative pubertal timing become less linked to age at menarche and more linked to race among older girls. Model 6 added interactions between age and schoolmates' pubertal timing, none of which was significant. Across all models, R^2 values ranged from .14 to .16, suggesting that friend- and school-level characteristics did not explain more variance than individual level predictors, and that most variation in relative pubertal timing was left unexplained, even by other self-report measures of pubertal timing.

Table 11. Predicting Male Relative Pubertal Timing from Individual, Friend, and School Variables

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Individual Variables						
Stage-Normative Timing	.42 (.02)***	.41 (.03)***	.39 (.03)***	.37 (.03)***	.38 (.03)***	.37 (.03)***
Age	.56 (.28)*	.63 (.36)	.80 (.41)*	.79 (.40)	.80 (.41)*	.86 (.42)*
Age-Squared	-.02 (.01)*	-.02* (.01)	-.03 (.01)*	-.03 (.01)*	-.03 (.01)*	-.03 (.01)*
Black	.06 (.07)	.04 (.03)	.05 (.17)	.05 (.17)	.04 (.18)	.04 (.18)
Hispanic	-.09 (.06)	-.07 (.11)	0.00 (.12)	-0.01 (.12)	-0.02 (.12)	-0.02 (.12)
Friend Variables						
Friends' Stage-Normative Timing		-.03 (.04)	-.06 (.04)	-.06 (.04)	-.08 (.05)	-.09 (.05)
Friends' Relative Timing		.05 (.04)	.04 (.04)	.04 (.04)	.03 (.04)	.03 (.04)
Friends' Mean Age		.09 (.04)*	.12 (.05)*	.11 (.05)*	.11 (.05)*	.12 (.05)*
% Black Friends		0.04 (.20)	0.16 (.21)	0.16 (.21)	0.16 (.21)	0.16 (.22)
% Hispanic Friends		0.00 (.14)	0.09 (.15)	0.09 (.15)	0.07 (.15)	0.08 (.15)
School Variables						
School Stage-Normative Timing			-.41 (.35)	-.35 (.35)	-.31 (.35)	-.34 (.36)
School Relative Timing			.85 (.37)*	.86 (.37)*	.91 (.37)*	.92 (.38)*
School age			-.04 (.06)	-.04 (.06)	-.04 (.06)	-.05 (.06)
% Black Students			0.00 (.00)	0.00 (.00)	0.00 (.00)	0.00 (.00)
% Hispanic Students			0.00 (.01)	0.00 (.01)	0.00 (.01)	0.00 (.01)
Moderation by Age						
Stage x Age				.03 (.02)	.03 (.02)	.03 (.02)
Friends' Stage x Age					.04 (.03)	.04 (.03)
Friends' Relative x Age					.02 (.02)	.02 (.02)
School Stage x Age						-.11 (.16)
School Relative x Age						.06 (.20)
Model R^2	.14	.15	.16	.16	.16	.16

*$p < .05$ **$p < .01$ ***$p < .001$

Table 12. Predicting Female Relative Pubertal Timing from Individual, Friend, and School Variables.

	Model 1	Model 2	Model 3	Model 4	Model 4a	Model 5	Model 6
Individual Variables							
Stage-Normative Timing	.39(.02)***	.39(.02)***	.39(.03)***	.37(.03)***	.37(.03)***	.37(.03)***	.36 (.03)***
Age at Menarche	-.15 (.02)***	-.16 (.02)***	-.15 (.02)***	-.16 (.02)***	-.17 (.02)***	-.17 (.02)***	-.17 (.02)***
Age	1.00 (.27)***	.97 (.34)***	1.04 (.42)*	1.15 (.42)**	.81 (.45)**	1.17 (.51)*	1.21 (.55)*
Age-Squared	-.03 (.01)***	-.03 (.01)**	-.03 (.01)**	-.04 (.01)**	-.04 (.01)**	-.03 (.01)**	-.04 (.01)*
Black	-.27 (.06)***	-.27 (.06)**	-.12 (.17)	-.10 (.17)	.08 (.18)	.09 (.18)	.10 (.18)
Hispanic	-.03 (.06)	-.08 (.11)	-.03 (.12)	-.03 (.12)	.03 (.12)	.03 (.12)	.03 (.12)
Friend Variables							
Friends' Stage-Normative Timing		.01 (.04)	-.02 (.04)	-.02 (.04)	-.02 (.04)	-.02 (.04)	-.02 (.04)
Friends' Relative Timing		.03 (.04)	.01 (.04)	-.01 (.04)	-.02 (.04)	-.02 (.04)	-.02 (.03)
Friends' Menarche		-.01 (.03)	-.02 (.03)	-.02 (.03)	-.01 (.03)	-.01 (.03)	-.01 (.03)
Friends' Mean Age		0.00 (.04)	.03 (.04)	.03 (.04)	.02 (.04)	.03 (.04)	.03 (.05)
% Black Friends		.09 (.19)	.13 (.21)	.10 (.21)	.05 (.21)	.05 (.21)	.06 (.21)
% Hispanic Friends		.07 (.14)	.26 (.14)	.26 (.14)	.27 (.14)	.28 (.14)	.29 (.15)*
School Variables							
School Stage-Normative Timing			-.02 (.37)	-.01 (.37)	-.17 (.37)	-.19 (.37)	-.22 (.38)
School Relative Timing			.66 (.43)	.62 (.44)	.52 (.44)	.52 (.44)	.47 (.45)
School Age			-.02 (.07)	-.03 (.07)	0.00 (.07)	0.00 (.07)	0.01 (.07)
School Menarche			.13 (.27)	.16 (.27)	.04 (.27)	.00 (.27)	-.04 (.29)
% Black Students			0.00 (.00)	0.00 (.00)	-.01 (.00)	-.01 (.00)	-.01 (.00)
% Hispanic Students			0.00 (.01)	0.00 (.01)	0.00 (.01)	0.00 (.01)	0.00 (.01)
Moderation by Age							
Stage x Age				-.01 (.02)	-.01 (.02)	-.01 (.02)	.00 (.02)
Menarche x Age				.03 (.01)*	.03 (.01)	.03 (.01)*	.03 (.01)*
Black x Age					-.22 (.08)**	-.23 (.08)**	-.24 (.08)**
Hispanic x Age					-.07 (.06)	-.08 (.06)	-.09 (.06)
Friends' Stage x Age						0.00 (.02)	0.00 (.02)
Friends' Relative x Age						0.00 (.02)	.01 (.02)
Friends' Menarche x Age						-.03 (.02)	-.03 (.02)
School Stage x Age							.03 (.17)
School Relative x Age							-.10 (.16)

Table 12, cont.

	Model 1	Model 2	Model 3	Model 4	Model 4a	Model 5	Model 6
School Menarche x Age							-.06 (.14)
Model R^2	.21	.22	.22	.22	.23	.23	.23

Table 13. Predicting Male Stage-Normative Pubertal Timing from Individual, Friend, and School Variables.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Individual Variables						
Relative Timing	.32 (.02)***	.32 (.02)***	.29 (.02)***	.29 (.03)***	.29 (.03)***	.29 (.03)***
Age	.08 (.24)	-.06 (.31)	.21 (.35)	.22 (.35)	.22 (.35)	.27 (.36)
Age-Squared	.00 (.01)	0.00 (.01)	-.01 (.01)	-.01 (.01)	-.01 (.01)	-.01 (.01)
Black	.32 (.06)***	.29 (.14)*	-.09 (.15)	-.09 (.15)	-.09 (.15)	-.08 (.15)
Hispanic	.12 (.05)*	.13 (.10)	-.08 (.10)	-.08 (.10)	-.08 (.10)	-.08 (.10)
Friend Variables						
Friends' Stage-Normative Timing		.17 (.03)***	.10 (.04)**	.10 (.04)**	.09 (.04)*	.09 (.04)*
Friends' Relative Timing		0.00 (.03)	-.01 (.03)	-.01 (.03)	-.01 (.03)	-.01 (.03)
Friends' Mean Age		.02 (.04)	.06 (.04)	.06 (.04)	.06 (.04)	.06 (.04)
% Black Friends		-.01 (.17)	.06 (.18)	.06 (.18)	.06 (.19)	.06 (.19)
% Hispanic Friends		-.02 (.12)	.28 (.13)*	.28 (.13)*	.28 (.13)*	.28 (.13)*
School Variables						
School Stage-Normative Timing			1.07 (.30)***	1.07 (.30)***	1.08 (.30)***	1.08 (.31)***
School Relative Timing			-.48 (.32)	-.48 (.32)	-.47 (.32)	-.51 (.33)
School age			-.09 (.05)	-.09 (.05)	-.09 (.05)	-.09 (.05)
% Black Students			0.00 (.00)	0.00 (.00)	0.00 (.00)	0.00 (.00)
% Hispanic Students			0.00 (.00)	0.00 (.00)	0.00 (.00)	0.00 (.00)
Moderation by Age						
Relative x Age				0.00 (.02)	0.00 (.02)	0.00 (.02)
Friends' Stage x Age					.01 (.02)	.01 (.02)
Friends' Relative x Age					0.00 (.02)	0.00 (.02)
School Stage x Age						-.01 (.14)
School Relative x Age						-.05 (.18)

*$p < .05$ **$p < .01$ ***$p < .001$

Table 14. Predicting Female Stage-Normative Pubertal Timing from Individual, Friend, and School Variables.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Individual Variables						
Relative Timing	.33 (.02)***	.33 (.02)***	.32 (.03)***	.33 (.03)***	.32 (.03)***	.32 (.03)***
Age at Menarche	-.03 (.02)	-.01 (.02)	-.01 (.02)	-.03 (.02)	-.03 (.02)	-.03 (.02)
Age	.14 (.25)	-.02 (.32)	-.49 (.39)	-.38 (.39)	-.40 (.39)	-.39 (.45)
Age-Squared	0.00 (.01)	0.00 (.00)	.01 (.01)	.01 (.01)	.01 (.01)	.01 (.01)
Black	.15 (.06)*	.25 (.14)	.28 (.16)	.29 (.16)	.29 (.16)*	.31 (.16)*
Hispanic	-.08 (.06)	.04 (.10)	.08 (.11)	.07 (.11)	.08 (.11)	.07 (.11)
Friend Variables						
Friends' Stage-Normative Timing		.01 (.04)	-.02 (.04)	-.02 (.04)	-.04 (.04)	-.03 (.04)
Friends' Relative Timing		.04 (.03)	.01 (.03)	.02 (.03)	.01 (.03)	.02 (.04)
Friends' Menarche		-.03 (.03)	-.05 (.03)	-.05 (.03)	-.06 (.03)*	-.07 (.03)*
Friends' Mean Age		.07 (.04)	.05 (.04)	.04 (.04)	.04 (.04)	.05 (.04)
% Black Friends		0.00 (.00)	.08 (.19)	.06 (.19)	.05 (.19)	.02 (.19)
% Hispanic Friends		-.15 (.13)	-.03 (.13)	-.03 (.13)	-.04 (.13)	-.04 (.13)
School Variables						
School Stage-Normative Timing			.98 (.34)**	.99 (.34)**	1.03 (.34)**	1.13 (.35)**
School Relative Timing			-.16 (.40)	-.19 (.40)	-.16 (.40)	-.01 (.42)
School Age			.05 (.06)	.04 (.06)	.04 (.06)	.09 (.07)
School Menarche			.01 (.25)	.02 (.25)	.08 (.25)	-.01 (.26)
% Black Students			0.00 (.00)	0.00 (.00)	0.00 (.00)	0.00 (.00)
% Hispanic Students			0.00 (.00)	0.00 (.00)	0.00 (.01)	0.00 (.01)
Moderation by Age						
Relative x Age				-.02 (.02)	-.02 (.02)	-.02 (.02)
Menarche x Age				.02 (.01)	.02 (.01)	.02 (.01)
Friends' Stage x Age					.03 (.02)	.01 (.02)
Friends' Relative x Age					.04 (.02)	.04 (.02)
Friends' Menarche x Age					.03 (.02)	.03 (.02)
School Stage x Age						.34 (.16)*
School Relative x Age						-.17 (.14)
School Menarche x Age						-.17 (.13)

*$p < .05$ **$p < .01$ ***$p < .001$

Do any measures of peers' pubertal timing predict individual stage-normative pubertal timing?

Results of regression analyses using stage-normative pubertal timing as the dependent variable are shown in Table 13 for males. Model 1 showed the expected positive association between self-reported relative and stage-normative timing ($\beta = .32$, $p < .001$). Neither age nor age-squared were directly associated with stage-normative timing, which was expected since stage-normative timing was itself age-standardized.

Model 2 added main effects of the adolescent's nominated friends. Consistent with initial correlational analyses, boys' stage-normative timing was positively associated with their male friends' stage-normative timing. There were no associations between boys' stage-normative timing and their friends' relative pubertal timing. There were no significant effects of friends' age, race, or ethnicity on males' stage-normative timing. Model 3 added school-level predictors of stage-normative pubertal timing. At the school level, there was a positive association between stage-normative timing and the average stage-normative timing reported by one's male schoolmates ($\beta = 1.07$, $p < .001$), suggesting school-level similarity for stage-normative timing. This result aligned with the school-level ICC estimated in our preliminary correlational analyses. The racial and ethnic composition of the school was not associated with boys' stage-normative pubertal timing. Subsequent models added interactions between age and individual (Model 4), friend (Model 5) and school (Model 6) characteristics. We found no significant interactions between age and any of these characteristics in predicting boys' stage-normative pubertal timing.

Parallel analyses for females are shown in Table 14. Model 1, which included individual characteristics only, showed positive associations between stage-normative and relative pubertal timing ($\beta = .33$, $p < .001$). Age at menarche was surprisingly not

significantly predictive of stage-normative timing. Black females had earlier stage-normative timing than their same-aged, White female peers ($\beta = .15, p < .05$).

Models 2 and 3 added characteristics of female friends and schoolmates, respectively. There was no evidence for an association between stage-normative timing and any measure of friends' pubertal timing. However, consistent with findings for boys, at the school level, there was a significant association between school and individual stage-normative timing ($\beta = .98, p < .01$). There were no friend-level or school-level effects of age, race, or ethnicity on females' self-reported stage-normative pubertal timing. The individual-level effect of Black race became non-significant when friend-level effects were added.

Model 4 examined interactions between age and age at menarche and between age and relative pubertal timing in predicting stage-normative timing. No age interactions were significant. Models 5 and 6 added interactions between age and friends' pubertal timing and schoolmates' pubertal timing, none of which were significant. The main effect of Black race was significant in these models, with Black girls reporting earlier stage-normative timing ($\beta = .03, p < .05$). In addition, a significant main effect of friends' age at menarche was found, such that girls whose friends had earlier ages at menarche reported earlier stage-normative pubertal timing.

Do race and ethnicity moderate associations between peer stage-normative timing and individual relative timing?

We conducted exploratory analyses to test whether the race and ethnicity moderated the hypothesized effects (i.e., an effect of peer stage-normative timing on individual relative pubertal timing). A series of models, analogous to the age moderation models, were estimated to test for interactions between Black race and peer pubertal timing (relative, stage-normative, and age at menarche) and between Hispanic ethnicity

and peer pubertal timing. These models revealed no significant moderating effects of race ethnicity. Full results from race/ethnicity moderation models are available from the first author upon request.

DISCUSSION

Self-report measures of pubertal timing correlate modestly with each other and with clinically assessed methods and are therefore often considered imperfect proxies for underlying biological changes (reviewed in Shirtcliff, Dahl, & Pollack, 2009; see also Dorn, Wald, Woodward, & Biro, 2006). Nevertheless, these measures predict an array of important health risks and behaviors in adolescents, indicating that perceptions of timing are meaningful psychological constructs. Moreover, different self-report measures often predict different outcomes (see Baams et al., 2015). While it has been theorized that accuracy of self-report measures may hinge on social comparisons (Moore, Harden, & Mendle, 2014), the relation between social contexts and self-perceptions of puberty has not yet been empirically tested. The current study addressed this gap in the literature by examining associations between self-reports of pubertal timing among adolescents, their friends, and their schoolmates. Our findings provide further support for a high degree of variability across self-report pubertal timing measures and suggest that these differences could be due, in part, to the peer context.

Three self-report measures of pubertal timing were examined in a high school sample of male and female adolescents: "stage-normative timing" (age-standardized ratings of body changes), "relative pubertal timing" (perceived development compared to one's same aged peers), and, in females, self-reported age at menarche. Initial analyses assessed how similar adolescents were to their nominated friends and to their schoolmates in these three measures, controlling for age, race, and ethnicity. Males

appeared similar to their nominated friends and schoolmates in stage-normative pubertal timing, whereas females were similar to their friends in relative pubertal timing. It may also be due to racial or ethnic differences in the social desirability of early vs. late pubertal timing.

Homophily by pubertal timing may reflect a selection process, with adolescents selecting friends who are similar to them in pubertal timing or in the behaviors that are correlated with pubertal timing. Such behaviors include athletic involvement (Malina & Bielicki, 1996), academic achievement, and various forms of risky behavior, including substance use and delinquency (reviewed in Ferrero & Mendle, 2012; Mendle, et al., 2007). It is well established that these behaviors cluster within adolescent peer groups (e.g., Matsueda & Anderson, 1998; Curran, Stice, & Chassin, 1997; Tolson & Urberg, 1993), and peer group selection is a commonly proposed mechanism for associations between pubertal timing and risky behavior (Caspi, Lynam, Moffitt, & Silva, 1993). Peer similarity for perceptions of development may also reflect an assimilation process, whereby adolescents perceive themselves as early maturing partly because their friends also appear older. For example, a typically developing girl (i.e., one with average objective pubertal timing) begins to affiliate with a risky, precocious peer group (e.g., girls who use drugs, have sex, wear makeup, and dress to look older). Regardless of her own age at menarche and her actual physical maturity, she may perceive herself to be more advanced in her pubertal development– because she may in fact look and act older – than her same-aged peers.

Differences in peer homophily across gender lend clues to how adolescents arrive at their self-conceptions. Females appeared similar to their friends on relative pubertal timing, rated on a scale ranging from "I look younger than most [girls my age]" to "I look older than most [girls my age]." Males appeared similar to their friends in stage-

normative timing, rated on a scale ranging from "how [you looked] in grade school" to how they imagine a "grown man" looks. Thus, males seemed alike in whether they view themselves as boys versus men, whereas females seemed alike in whether they see themselves as typical versus atypical. In addition, among males, the age of nominated friends predicted self-reported relative pubertal timing: having older friends was correlated with earlier relative pubertal timing. This finding is consistent with previous studies that suggests early maturing males tend to affiliate with older peers (Ge et al., 2002), and, as a result, engage in risky behavior at an earlier age than their typically developing peers (Mendle & Ferrero, 2012).

Our study further examined associations between relative pubertal timing and peers' stage-normative timing. Controlling for an individual's self-reported physical changes (breast growth, voice changes, menarche, body hair – all factors that should theoretically predict a global, single item rating of pubertal timing), we tested whether the pubertal timing of one's nominated friends and schoolmates influenced self-reported relative pubertal timing. However, we found no evidence for peer contrast effects, and most of the variance in relative pubertal timing (in both males and females) remained unexplained in the models tested.

There are several possible explanations for this null finding. We operationalized "stage-normative peer pubertal timing" as friends' and schoolmates' ratings of their own body changes. A peer contrast effect is basically a negative bias—a distorted perception of one's actual pubertal timing due to the actual, observable reference point being skewed. For the purposes of these analyses we assume that the peers' reports of stage-normative and peer normative timing indexed something that is stable and accurate enough for an individual to use as a reference group. We cannot, however, confirm this. In addition, the adolescents in this sample were, for the most part, post-pubertal or mid-

pubertal. It is possible that adolescents are more attentive to – and thus influenced by -- perceptions of their friends when they are in the midst of pubertal changes and when there is more variability in pubertal status. Moreover, although adolescents spend a great deal of time in school and with friends, there are numerous other targets of comparison to which adolescents are regularly exposed, such as family members or figures in the media, who may shape their perceptions of maturation. Finally, the current study used a school-wide nonclinical sample. It is possible that peer effects on perceived pubertal timing would emerge in more specialized contexts that attract adolescents with off-time development, such as athletic environments. In sum, the current study does not support the hypothesis that peers' stage-normative pubertal timing influences perceptions of one's relative pubertal timing, but a more objective rating of peer pubertal timing and a different study sample might yield different findings.

Comparing across the different models of stage-normative and relative pubertal timing, an interesting pattern of racial difference emerges. Although Black girls and boys reported earlier (more advanced) pubertal timing than White adolescents on the stage-normative measure, this racial difference was not reflected in self-perceptions of relative pubertal timing. There was no significant racial difference in relative pubertal timing for boys, and Black girls actually perceived themselves as later developing, on average, than White girls perceived themselves to be. One possible interpretation for these findings is that, although Black and Hispanic females develop earlier than their White counterparts, they do not view themselves as early maturing. This could indeed be due to a social comparison process, with adolescents using their same-race peers as the target of comparison; this is why the racial composition of the peer group is included as a variable in our regression models.

Results from the current study must be interpreted in light of several limitations. First the data are cross-sectional, and do not explain within-person changes in perceived development over time. There are also limitations to the measures of pubertal timing used. The reliability of reports of age at menarche decreased as Add Health participants aged. Relative pubertal timing was also assessed using a single question; although consistent with previous research practices (Dubas et al., 1991; Graber et al., 1997), a more extensive assessment of perceived development might permit a better understanding of how individuals leverage peer comparisons to interpret their development.

Other limitations pertain to the study sample. The sample is from 1994-1995 and therefore not representative of today's youth, potentially limiting the generalizability of the findings. The sample was moreover limited to adolescents with identifiable (same school) peer nominations, and this sample differs slightly from the full sample, as shown in Table 1. Previous studies of the Wave I Add Health social network data have revealed some differences between adolescents with and without identifiable friends. Female adolescents without identifiable friends were more likely to be Hispanic (Kretsch, Mendle, & Harden, 2014) and there appeared to be slightly lower rates of alcohol use among adolescents without identifiable peer nominations (Cruz, Emery, & Turkheimer, 2012). However, since the focus of this study was how characteristics of the friendship reference group might shape perceptions of one's own pubertal timing, it was reasonable to limit our sample to individuals who have such a reference group.

The measure of friends' timing was also restricted to same-sex friends. We focused on same-sex friends because the measure of relative timing asked adolescents to compare themselves to other same-sex peers. Given the increased frequency of opposite-sex friendships in adolescence, future studies might explore the associations between gender composition of the peer group and the pubertal timing of one's opposite-sex peers

as correlates of perceived pubertal timing. Regarding gender differences, we did not test statistically test gender as a moderator of the effect of peers' pubertal timing on individual timing. Rather, we analyzed boys and girls data separately and examined if results of our main research questions seemed consistent across male and female populations. The gender differences we did identify, particularly regarding school-level clustering of relative and stage-normative pubertal timing, warrant further investigation.

Identity development is a key task of adolescence, and part of identity development is negotiating one's place in the social world. Peer comparisons are ubiquitous in adolescence and measures of pubertal timing implicitly or explicitly invoke a peer comparison. Therefore, the pubertal timing of one's peer group may explain variation in adolescents' subjective reports of pubertal timing. The current study found evidence for peer group similarity in perceived pubertal timing. Peer selection and socialization shape development, so this homophily may result from adolescents selecting peers who look similar to them in terms of physical development or because adolescents come to see themselves and present themselves as early maturing if they are in an early maturing group. Males were more likely to rate themselves as similar to their peer group on stage-normative timing, whereas females were more likely to rate themselves as similar to peers in subjective development. These results provide further evidence that relative pubertal timing, stage-normative timing, and age at menarche are unique indicators of the enduring impressions of adolescents' experience of puberty.

Chapter 5: Synthesis

The overall goal of this project was to explore the associations between puberty and adolescent risk-taking. The three studies focused on individual differences—specifically, how variation in the timing, context, and perception of this universal milestone might contribute to individual differences in risky behavior. Table 15 summarizes the methods and key findings of each study. This set of studies began by testing a basic biological pathway from genes to hormones to behavior, and moved toward increasingly abstract and subjective questions about perceptions and self-comparisons. In closing, I will revisit some of the theories and themes that informed this project in light of current study findings, and I will comment on implications for future research.

RAGING HORMONES

Study 1 examined a biological mechanism for the association between pubertal timing and risk-taking. Combining self-report, hormonal, and behavioral measures, this study used a twin design to test the hypothesis that testosterone mediated genetic risk for substance use via its effect on reward seeking. The primary hypothesis was not supported. Although testosterone was heritable in both males and females, there were no phenotypic association between testosterone and initiation of substance use, neither directly nor through reward seeking. Self-reported pubertal timing, which was correlated with testosterone in males, was also unrelated to reward seeking. A number of explanations for these unexpected findings were considered, including the age and level of substance use of the sample, the measurement of reward-seeking, and unmeasured genetic, contextual, and biological variables that may moderate the effects of pubertal timing on substance use initiation.

PUBERTY IN CONTEXT

Study 2 examined the possibility that social context – in particular, peer groups -- moderated the association between risk-taking and pubertal timing. In this study, the risk-taking phenotype was delinquency among adolescent girls. A twin design was used to test whether pubertal timing moderated the quasi-causal (i.e., within-twin-pair) association between peer delinquency and individual delinquency. Consistent with the "contextual amplification" hypothesis, earlier pubertal timing magnified the quasi-causal association between peer and individual risk-taking. Study 2 added a contextual layer that was missing from Study 1, by combining data from adolescent girls, their friends, and their sisters. Results illustrated how behavioral correlates of pubertal timing depend on the peer context.

Cross-sectional, non-experimental designs leave ambiguities regarding the relative roles of selection versus socialization in explaining the similarity among peers. Study 2 used a quasi-experimental, co-twin-control approach, and found that girls with earlier pubertal timing were more similar to their peers. This within-twin-pair association leaves open the possibility that earlier maturing girls were greater sources, rather than targets, of influence. Experimental studies that manipulate context can shed some light on directions of causal influence.

Study 2 focused on one aspect of social context, the peer environment, but future research should consider other aspects of the macro-environment. The cultural context surrounding puberty is transforming, and it is important to consider how risk-taking may manifest differently as contexts change. For example, the Internet provides new forums for risk-taking, novelty-seeking, social interaction, and also is a source of education about

puberty. Topics of increasing interest to researchers include the association between pubertal timing and exposure to Internet pornography (e.g., Skoog, Stattin, & Kerr, 2009), use of the Internet to access information about puberty and sexual health, associations between social network activity and pubertal timing (Skoog, Sorbring, & Bohlin, 2015), and the connections between Internet usage and risky sexual behavior (Braun-Caurville, Rojas, 2009).

DEFINING AND MEASURING PUBERTY

This set of studies relied primarily on self-report measures of pubertal development, which have been criticized as imprecise measures of actual physical development. Study 1 included a single measure of salivary testosterone. Testosterone was correlated with self-reported pubertal status in males, suggesting that these two measures were capturing variability in the same underlying process. However, this hormonal variability did not predict outcomes in Study 1. Studies 2 and 3 distinguished between three self-report measures of pubertal timing, and findings differed across these measures. Study 2 included three measures of pubertal timing: age at menarche, age-standardized ratings of body changes (similar to the self-report measure of pubertal timing in Study 1), and a peer comparison measure, which was a single question that asked girls to compare their physical development to their peers. There was no interaction between peer delinquency and age-standardized ratings of body changes, but there was an interaction between peer delinquency and the peer comparison measure. Perceived relative pubertal timing is not perfectly correlated with other measures, and continues to vary beyond the age at which the physical changes are generally complete. If variation in this construct is relevant for understanding risk-taking, understanding what contributes to

this variation is an important undertaking. In other words, what predicts perceived pubertal timing beyond actual pubertal timing?

Study 3 addressed this question by testing whether pubertal timing reported by one's friends and schoolmates related to perceived pubertal timing. Rather than focusing on risk-taking itself as an outcome variable, Study 3 used social network data to test how peer and school contexts shaped adolescents' perceptions of pubertal timing. Results were somewhat surprising. There was no evidence for a peer contrast or bias effect, but there was evidence that subjective measures of pubertal timing clustered within friendship groups. There were also notable gender differences. Males appeared similar to their friends in whether they viewed themselves as boys versus men, whereas girls appeared similar to friends in whether they perceived themselves as early versus late. This study, notable because it was the first to examine peer group similarity for perceived pubertal timing, left several important questions unanswered, such as whether similarity among peers changes over time and how peer similarity in pubertal timing may influence the individual experience of puberty and its effects on health outcomes, including risk-taking behaviors.

Collectively, these studies add to a growing body of evidence that self-report measures are psychologically meaningful (reviewed in Mendle, 2014). Future research would benefit from deliberate distinction between biological and subjective measures of development. Studies using self-report measures of pubertal development should note that they are measuring children's interpretations of physical changes, and should interpret findings accordingly. As noted in a recent review, "children are not trained medical professionals, and their self-reports represent their own truths" (Mendle, 2014; p. 217). As we try to understand the impact of puberty, we should not only acknowledge but also take advantage of this distinction.

MOVING FORWARD: THE WIDENING MATURITY GAP

As discussed in the Introduction, many explanations for the link between pubertal timing and health outcomes emphasize maturity gaps. Cultural, social, emotional, and cognitive development follow unique developmental trajectories, and these maturity gaps are thought to explain both group differences and individual differences in health risk behavior. These maturity gaps are widening, as the age at which puberty begins has declined for both boys and girls (Lee & Styne, 2013). When Marshall and Tanner (1969; 1970) first published on the Tanner Stages, the average age of breast development was 11.15 years; in the 1990s it was 9 years (earlier for African American girls); and by 2010, one-third of girls begin developing breasts at 7 or 8 (Biro 2010). Assessing the secular trend in male pubertal development has been more difficult, but there is some evidence that boys are beginning puberty approximately 1.5 years earlier than they did half a century ago (Herman-Giddens et al., 2012; Marshall & Tanner, 1970). The "secular trend" in physical development is of concern to parents, scientists, and educators, resulting in a growing awareness of the potential adverse outcomes of earlier pubertal timing (e.g., Greenspan & Deardorff, 2014).

At the same time, the age at which children reach "social maturity" in the U.S. continues to increase. The third decade in life is now considered "emerging adulthood," with numerous role transitions -- including marriage, parenthood, entry into the workforce, and financial independence from parents – occurring later and in more variable order (Arnett, 2000). If the disparity between biological and psychosocial maturity contributes to higher rates of risk-taking in adolescence, the widening maturity gap may not bode well for today's youth.

However, despite the widening maturity gap, rates of risk-taking have not increased. Substance use among high schoolers has actually declined (Johnston et al.,

2014), as have rates of delinquency (Sickmund & Puzzanchera, 2014). Much of the research linking pubertal timing with health outcomes is several decades old and warrants replication in updated sample. An important question for future research, therefore, is whether individual differences in puberty will continue to matter in the same way for adolescents. Moreover, as many authors have noted, perceptions of being "on time" or "early" are often incongruent with objective ratings (Dorn & Biro, 2011), yet these perceptions seem to matter a great deal for psychological outcomes (reviewed in Mendle, 2014). Will the psychological significance of puberty change if it occurs earlier for all adolescents? To what extent do individual differences at puberty explain variation in adolescent risk-taking? Overall, this research illustrates the complexities inherent in measuring puberty and highlights the role of context and perception in understanding its impact.

Table 15. Methods, findings, and implications of Studies 1, 2, and 3

	Study 1	Study 2	Study 3
Sample / Cohort	• Texas Twin Project • N = 530 • 50% female • Ages 13-20 • 2011-2013	• Add Health • N = 496 • 100% female • Ages 13-20 • 1994-1995	• Add Health • N = 2817 • 50% female • Ages 13-20 • 1994-1995
Measures of puberty	• Salivary testosterone, standardized by age, age^2, within gender • Pubertal development scale	• Age at menarche • Peer normative timing • Stage normative timing	• Age at menarche • Peer normative timing • Stage normative timing
Measures of risk-taking	• "Have you ever had a drink containing alcohol?" • "How often in the past year have you gotten drunk or high on alcohol?" • "Have you ever smoked marijuana?"	• "How often in the past year have you… • smoked cigarettes?" • consumed alcohol?" • gotten drunk?" • raced on a bike or skateboard or in a car or boat?" • lied to your parents?" • skipped school?" • done something dangerous because you were dared to?"	
Measures of context	• SES, race, ethnicity • Shared environment	• Shared environment • Peer group characteristics	• Race, ethnicity • Peer group characteristics • School characteristics
Sources of data	• Biological • Behavioral • Self report	• Self report • Nominated friends' report	• Self report • Nominated friends' report • Schoolmates' report
Analytic design	• Cross sectional • Latent factors measurement model • Univariate quantitative • genetic models	• Cross sectional • Sibling comparison • Multivariate quantitative genetic models	• Cross sectional • Intraclass correlations • Multiple linear regression

Table 15, cont.

	Study 1	Study 2	Study 3
Notable findings	• No phenotypic association between testosterone, reward seeking, and substance use initiation • Most behavioral measures loaded on a single factor, substantial task-specific variance, minimal overlap of behavioral and self-report • Strongest predictors of substance use were age, race, cognitive ability, and self-reported impulsivity	• Girls with earlier ages at menarche and earlier "peer normative timing" were more similar to their nominated friends in delinquency • The within-pair effect of having delinquent friends was stronger for early maturing girls • Supports contextual amplification hypothesis	• No evidence for a peer bias or peer contrast effect • Females were similar to their peers in peer-normative or relative pubertal timing • Males were similar to their peers in stage-normative timing • School-level similarity for males' stage-normative timing
Limitations and future directions	• Cross sectional • Low prevalence of substance use • Single hormone at single time point • Post/mid pubertal sample • No objective measures of pubertal timing	• Cross sectional • Self-report • Low level of delinquency • Post/mid pubertal sample • No objective measures of pubertal timing	• No objective measures of pubertal timing • Self-report • Missing data • Post/mid pubertal

References

Adams, G. R., & Marshall, S. K. (1996). A developmental social psychology of identity: Understanding the person-in-context. *Journal of Adolescence*, *19*(5), 429–442.

Alexander, G. M., Packard, M. G., Hines, M. (1994). Testosterone has rewarding affective properties in male rats: Implications for the biological basis of sexual motivation. *Behavioral Neuroscience, 108(2)*, 424-428.

Allen, J. P., Porter, M. R., & McFarland, C. F. (2006). Leaders and followers in adolescent close friendships: Susceptibility to peer influence as a predictor of risky behavior, friendship instability, and depression. *Development and Psychopathology, 18,* 155-172.

Andersson, T., & Magnusson, D. (1990). Biological maturation in adolescence and the development of drinking habits and alcohol abuse among young males: A prospective longitudinal study. *Journal of Youth and Adolescence, 19,* 33–41.

Angold, A., & Costello, E. J. (2006). Puberty and depression. *Child and Adolescent Psychiatric Clinics of North America*, *15*, 919-937.

Anokhin, A. P., Golosheykin, S., Grant, J., & Heath, A. C. (2009). Heritability of risk-taking in adolescence: a longitudinal twin study. *Twin research and human genetics: the official journal of the International Society for Twin Studies*, *12*(4), 366.

Apter, D. (1980). Serum steroids and pituitary hormones in female puberty: A partly longitudinal study. *Clinical Endocrinology, 12,* 107– 120.

Apter, S. J., & Eriksson, C. J. (2003). The effect of alcohol on testosterone concentrations in alcohol‐preferring and non‐preferring rat lines. *Alcoholism: Clinical and Experimental Research, 27,* 1190-1193.

Archer, J. (2006). Testosterone and human aggression: an evaluation of the challenge hypothesis. *Neuroscience & Biobehavioral Reviews*, *30*(3), 319-345.

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, *55*(5), 469.

Baker, J. H., Thornton, L. M., Lichtenstein, P., & Bulik, C. M. (2012). Pubertal development predicts eating behaviors in adolescence. *International Journal of Eating Disorders, 45(7)*, 819–826.

Beyth-Marom R, Austin L, Fischoff B, Palmgren C, Jacobs-Quadrel M. (1993). Perceived consequences of risky behaviors: Adults and adolescents. *Developmental Psychology, 29,* 549–563.

Biehl, M. C., Natsuaki, M. N., & Ge, X. (2007). The influence of pubertal timing on alcohol use and heavy drinking trajectories. *Journal of Youth and Adolescence, 36,* 153–167.

Biro, F. M., Lucky, A. W., Huster, G. A., & Morrison, J. A. (1995). Pubertal staging in boys. *Journal of Pediatrics, 127*(1), 100-102.

Blakemore, S. J., Burnett, S., & Dahl, R. E. (2010). The role of puberty in the developing adolescent brain. *Human Brain Mapping, 31*, 926-933.

Blanton, H. (2001). Evaluating the self in the context of another: The three-selves model of social comparison assimilation and contrast. In *Cognitive social psychology: The Princeton symposium on the legacy and future of social cognition* (pp. 75–87). Erlbaum Mahwah, NJ.

Blyth, D. A., Simmons, R. G., & Zakin, D. F. (1985). Satisfaction with body image for early adolescent females: The impact of pubertal timing within different school environments. *Journal of Youth and Adolescence*, *14*(3), 207-225.

Bonat, S., Pathomvanich, A., Keil, M. F., Field, A. E., & Yanovski, J. A. (2002). Self-assessment of pubertal stage in overweight children. *Pediatrics*, *110*(4), 743–747.

Braams, B. R., van Duijvenvoorde, A. C., Peper, J. S., & Crone, E. A. (2015). Longitudinal Changes in Adolescent Risk-Taking: A Comprehensive Study of Neural Responses to Rewards, Pubertal Development, and Risk-Taking Behavior. *The Journal of Neuroscience*, *35*(18), 7226-7238.

Braun-Courville, D. K., & Rojas, M. (2009). Exposure to sexually explicit web sites and adolescent sexual attitudes and behaviors. *Journal of Adolescent Health*, *45*(2), 156-162.

Brechwald, W. A., & Prinstein, M. J. (2011). Beyond homophily: A decade of advances in understanding peer influence processes. *Journal of Research on Adolescence*, *21*(1), 166–179.

Brooks-Gunn, J., & Warren, M. P. (1989). Biological and social contributions to negative affect in young adolescent girls. *Child Development*, 40-55.

Brooks-Gunn, J., Attie, I., Burrow, C., Rosso, J., & Warren, M. (1989). The impact of puberty on eating concerns in athletic and non-athletic contexts. *Journal of Early Adolescence. Special issue: Early adolescent transitions: longitudinal studies of biological, psychological, and social interactions, 9*, 269-290.

Buchanan, C. M., Eccles, J. S., & Becker, J. B. (1992). Are adolescents the victims of raging hormones? Evidence for activational effects of hormones on moods and behavior at adolescence. *Psychological bulletin*, *111*(1), 62.

Burt, S. A., McGue, M., DeMarte, J. A., Krueger, R. F., & Iacono, W. G. (2006). Timing of menarche and the origins of conduct disorder. *Archives of General Psychiatry, 63*, 890-896.

Campbell, B. C., Dreber, A., Apicella, C. L., Eisenberg, D. T., Gray, P. B., Little, A. C., … & Lum, J. K. (2010). Testosterone exposure, dopaminergic reward, and sensation seeking in young men. *Physiology & Behavior, 99*, 451-456.

Cance, J. D., Ennett, S. T., Morgan-Lopez, A. A., & Foshee, V. A. (2012). The stability of perceived pubertal timing across adolescence. *Journal of Youth and Adolescence*, *41*(6), 764–775.

Caramaschi, D., Booij, L., Petitclerc, A., Boivin, M., & Tremblay, R. E. (2012). Genetic and environmental contributions to saliva testosterone levels in male and female infant twins. Psychoneuroendocrinology, 37, 1954-1959.

Carter, R., Caldwell, C. H., Matusko, N., Antonucci, T., & Jackson, J. S. (2011). Ethnicity, perceived pubertal timing, externalizing behaviors, and depressive

symptoms among black adolescent girls. *Journal of Youth and Adolescence*, *40*(10), 1394–1406.

Carter, R., Silverman, W. K., & Jaccard, J. (2013). Race and perceived pubertal transition effects on girls' depressive symptoms and delinquent behavior. *Journal of Youth and Adolescence, 42,* 1155-1168.

Casey, B. J., Getz, S., & Galvan, A. (2008). The adolescent brain. *Developmental Review, 28,* 62–77.

Casey, V., Dwyer, J., Coleman, K., Krall, E., Gardner, J., & Valadian, I. (1991). Accuracy of recall by middle-aged participants in a longitudinal study of their body size and indices of maturation earlier in life. *Annals of Human Biology*, *18*(2), 155–166.

Caspi, A., Lynam, D., Moffitt, T. E., & Silva, P. A. (1993). Unraveling girls' delinquency: biological, dispositional, and contextual contributions to adolescent misbehavior. *Developmental Psychology*, *29*(1), 19–30.

Cauffman, E., Shulman, E., Steinberg, L. Claus, E., Banich, M., Graham, S., & Woolard, J. (2010). Age differences in affective decision making as indexed by performance on the Iowa Gambling Task. *Developmental Psychology, 46,* 193-207.

Cavanagh, S., Riegle-Crumb, C., & Crosnoe, R. (2007). Puberty and the Education of Girls. *Social Psychology Quarterly 70(2),* 186-98.

Cavanagh, S. (2004). The sexual debut of girls in early adolescence: The intersection of race, pubertal timing, and friendship group characteristics. *Journal of Research on Adolescence, 14,* 285-312.

Centers for Disease Control and Prevention (2009). Surveillance for violent deaths – National violent death reporting system, 16 States, 2006. Morbidity and Mortality Weekly Report: Surveillance Summaries, 58(SS-1). http://www.cdc.gov/mmwr/PDF/ss/ss5801.pdf

Charney, E. (2012). Behavior genetics and postgenomics. *Behavioral and Brain Sciences, 35,* 331-358.

Chassin, L., Hussong, A., & Beltran, I. (2004). Adolescent Substance Use. In R. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology,* (3rd el., pp. 723-763). New York: Wiley.

Chein, J., Albert, D., O'Brien, L., Uckert, K., & Steinberg, L. (2010). Peers increase adolescent risk taking by enhancing activity in the brain's reward circuitry. *Developmental Science, 14,* F1-F10.

Clark, A.S., Lindenfeld, R.C., Gibbons, C.H. (1996). Anabolic–androgenic steroids and brain reward. *Pharmacology Biochemistry and Behavior*, *53*(3), 741-745.

Cohen-Cole, E., & Fletcher, J. M. (2008). Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic. *Journal of Health Economics, 27(5)*, 1382–1387.

Conley, C. S., & Rudolph, K. D. (2009). The emerging sex difference in adolescent depression: Interacting contributions of puberty and peer stress. *Development and Psychopathology*, *21*(02), 593–620.

Costello, E. J., Sung, M., Worthman, C., & Angold, A. (2007). Pubertal maturation and the development of alcohol use and abuse. *Drug and Alcohol Dependence*, *88*, S50-S59.

Curran, P. J., Stice, E., & Chassin, L. (1997). The relation between adolescent alcohol use and peer alcohol use: A longitudinal random coefficients model. *Journal of Consulting and Clinical Psychology*, *65*(1), 130–140.

Cruz, J. E., Emery, R. E., & Turkheimer, E. (2012). Peer network drinking predicts increased alcohol use from adolescence to early adulthood after controlling for genetic and shared environmental selection. *Developmental Psychology, 48,* 1390.

Dabbs Jr, J. M., Jurkovic, G. J., & Frady, R. L. (1991). Salivary testosterone and cortisol among late adolescent male offenders. *Journal Of Abnormal Child Psychology*, *19*(4), 469-478.

Department of Justice (2003). Uniform Crime Report statistics. Federal Bureau of Investigation.

de Macks, Z. A., Moor, B. G., Overgaauw, S., Güroğlu, B., Dahl, R. E., & Crone, E. A. (2011). Testosterone levels correspond with increased ventral striatum activation in response to monetary rewards in adolescents. *Developmental Cognitive Neuroscience, 1,* 506-516.

de Water, E., Braams, B. R., Crone, E. A., & Peper, J. S. (2013). Pubertal maturation and sex steroids are related to alcohol use in adolescents. *Hormones and Behavior, 63,* 392–397.

Deppen, A., Jeannin, A., Michaud, P.-A., Alsaker, F., & Suris, J.-C. (2012). Subjective pubertal timing and health-compromising behaviours among Swiss adolescent girls reporting an on-time objective pubertal timing. Acta Paediatrica, 101(8), 868–872.

Dick, D. M., Rose, R. J., Viken, R. J., & Kaprio, J. (2000). Pubertal timing and substance use: Associations between and within families across late adolescence. *Developmental Psychology, 36,* 180-189.

Dishion, T. J., & Dodge, K. A. (2005). Peer contagion in interventions for children and adolescents: Moving towards an understanding of the ecology and dynamics of change. *Journal of Abnormal Child Psychology, 33*(3), 395–400.

Dorn, L. D., & Biro, F. M. (2011). Puberty and its measurement: A decade in review. *Journal of Research on Adolescence, 21,* 180–195.

Dorn, L., Dahl, R., Woodward, H., & Biro, F. M. (2006). Defining the boundaries of early adolescents: A user's guide to assessing pubertal status and pubertal timing in research with adolescents. *Applied Developmental Science*, *10*, 30–56.

Dorn, L. D., Sontag-Padilla, L. M., Pabst, S., Tissot, A., & Susman, E. J. (2013). Longitudinal reliability of self-reported age at menarche in adolescent girls: Variability across time and setting. *Developmental Psychology*, *49*(6), 1187.

Dubas, J. S., Graber, J. A., & Petersen, A. C. (1991). A longitudinal investigation of adolescents' changing perceptions of pubertal timing. *Developmental Psychology*, *27*(4), 580–86.

Duke, S. A., Balzer, B. W., & Steinbeck, K. S. (2014). Testosterone and its effects on human male adolescent mood and behavior: A systematic review. *Journal of Adolescent Health*, *55*(3), 315-322.

Dyk, P. H. (1993). Anatomy, physiology, and gender issues in adolescence. In T. P. Gullotta, G. R. Adams & R. Montemayor (Eds.), *Adolescent Sexuality* (pp. 35–56). Thousand Oaks, CA: Sage.

Edwards, D. A., & Casto, K. V. (2013). Women's intercollegiate athletic competition: Cortisol, testosterone, and the dual-hormone hypothesis as it relates to status among teammates. *Hormones and Behavior*, *64*(1), 153-160.

Eisenegger, C., Haushofer, J., Fehr, E. (2011). The role of testosterone in social interaction. *Trends in Cognitive Sciences, 15*, 263-271.

Erikson, E. H. (1993). *Childhood and society*. New York: WW Norton & Company.

Eriksson, C. P., Kaprio, J., Pulkkinen, L., & Rose, R. J. (2005). Testosterone and alcohol use among adolescent male twins: Testing between-family associations in within-family comparisons. *Behavior Genetics, 35,* 359-368.

Federal Bureau of Investigation (2012). *Crime in the United States, 2012*. Retrieved from http://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2012/crime-in-the-u.s.-2012

Fergusson, D. M., Vitaro, F., Wanner, B., & Brendgen, M. (2007). Protective and compensatory factors mitigating the influence of deviant friends on delinquent behaviours during early adolescence. *Journal of Adolescence, 30,* 33-50.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140.

Finkenauer, C., Engels, R. C., & Meeus, W. (2002). Keeping secrets from parents: Advantages and disadvantages of secrecy in adolescence. *Journal of Youth and Adolescence*, *31*(2), 123–136.

Forbes, E. E., Ryan, N. D., Phillips, M. L., Manuck, S. B., Worthman, C. M., Moyles, D. L., . . . Dahl, R. E. (2010). Healthy adolescents' neural response to reward: Associations with puberty, positive affect, and depressive symptoms. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*, 162–172.

Forbes, E.E., & Dahl, R.E. (2010). Pubertal development and behavior: Hormonal activation of social and motivational tendencies. *Brain and Cognition, 72,* 66-72.

Foster, H., Hagan, J., & Brooks-Gunn, J. (2008). Growing up fast: Stress exposure and subjective 'weathering' in emerging adulthood. *Journal of Health and Social Behavior, 49(2)*, 162 – 177.

Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: An experimental study. Developmental Psychology, 41, 625– 635.

Ge, X., Brody, G. H., Conger, R. D., Simons, R. L., & McBride- Murray, V. (2002). Contextual amplification of pubertal transition effects on deviant peer affiliation and externalizing behavior among African American children. *Developmental Psychology, 38*, 42–54.

Ge, X., Conger, R. D., & Elder, G. H. (1996). Coming of age too early: Pubertal influences on girls' vulnerability to psychological distress. *Child Development, 67,* 3386-3400.

Ge, X., Natsuaki, M. N., Neiderhiser, J. M., & Reiss, D. (2007). Genetic and environmental influences on pubertal timing: Results from two national sibling studies. *Journal of Research on Adolescence, 17*, 767-788.

Ge, X. & Natsuaki, M.N. (2009). In search of explanations for early pubertal timing effects on developmental psychopathology. *Current Directions in Psychological Science, 18,* 327-331.

Gottesman, I.I., & Gould, T.D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *American Journal of Psychiatry, 160*, 636-645.

Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford, CA: Stanford University Press.

Graber, J. A., Lewinsohn, P. M., Seeley, J. R., & Brooks-Gunn, J. (1997). Is psychopathology associated with the timing of pubertal development? *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 1768–1776.

Graber, J. A., Seeley, J. R., Brooks-Gunn, J., & Lewinsohn, P. M. (2004). Is pubertal timing associated with psychopathology in young adulthood? *Journal of the American Academy of Child & Adolescent Psychiatry*, *43*(6), 718–726.

Granger, D. A., Schwartz, E. B., Booth, A., & Arentz, M. (1999). Salivary testosterone determination in studies of child health and development. *Hormones and Behavior, 35*(1), 18-27.

Greenspan, L., & Deardorff, J. (2014). *The New Puberty: How to Navigate Early Development in Today's Girls*. Rodale.

Grumbach, M.M., Styne, D.M., 2003. Puberty ontogeny, neuroendocrinology, physiology, and disorders. In: Larsen, P.R., Kronenberg, H.M., Melmed, S., Polonski, K.S. (Eds.), Williams Textbook of Endocrinology. 10th ed. Elsevier, New York, pp. 1115—1286.

Halpern, C., Udry, R., & Suchindran, C. (1998). Monthly Measures of Salivary Testosterone Predict Sexual Activity in Adolescent Males. *Archives of Sexual Behavior, 27*(5), 445-465.

Han, C., McGue, M. K., & Iacono, W. G. (1999). Lifetime tobacco, alcohol and other substance use in adolescent Minnesota twins: univariate and multivariate behavioral genetic analyses. *Addiction*, *94*(7), 981-993.

Harden, K. P., Hill, J. E., Turkheimer, E., & Emery, R. E. (2008). Gene—environment correlation and interaction in peer effects on adolescent alcohol and tobacco use. *Behavior Genetics, 38,* 339-347.

Harden, K. P., & Mendle, J. (2012). Gene-environment interplay in the association between early pubertal timing and delinquency in adolescent girls. *Journal of Abnormal Psychology, 121,* 73-87.

Harden, K. P., Quinn, P. D., & Tucker-Drob, E. M. (2012). Genetically influenced change in sensation seeking drives the rise of delinquent behavior during adolescence. *Developmental Science, 15,* 150-163.

Harden, K. P., Kretsch, N., Tackett, J. L., & Tucker-Drob, E. M. (2014). Genetic and environmental influences on testosterone in adolescents: Evidence for sex differences. *Developmental Psychobiology*, 56(6), 1278-1289.

Haynie, D. L. (2003). Contexts of risk? Explaining the link between girls' pubertal timing and their delinquency involvement. *Social Forces, 82*, 355-397.

Haynie, D. L., & Piquero, A. R. (2006). Pubertal development and physical victimization in adolescence. *Journal of Research in Crime and Delinquency*, *43*(1), 3–35.

Heath, A. C., Kessler, R. C., Neale, M. C., Hewitt, J. K., Eaves, L. J., & Kendler, K. S. (1993). Testing hypotheses about direction of causation using cross-sectional family data. *Behavior Genetics, 23*(1), 29-50.

Herman-Giddens, M. E., Steffes, J., Harris, D., Slora, E., Hussey, M., Dowshen, S. A., ... & Reiter, E. O. (2012). Secondary sexual characteristics in boys: data from the Pediatric Research in Office Settings Network. *Pediatrics*, *130*(5), e1058-e1068.

Hoekstra, R. A., Bartels, M., & Boomsma, D. I. (2006). Heritability of testosterone levels in 12-year-old twins and its relation to pubertal development. *Twin Research and Human Genetics*, *9*(04), 558-565.

Hollenstein, T., & Lougheed, J. P. (2013). Beyond storm and stress: Typicality, transactions, timing, and temperament to account for adolescent change. *American Psychologist, 68*, 444-454.

Hopfer, C. J., Crowley, T. J., & Hewitt, J. K. (2003). Review of twin and adoption studies of adolescent substance use. *Journal of the American Academy of Child & Adolescent Psychiatry*, *42*(6), 710-719.

Huizink, A. C., Levälahti, E., Korhonen, T., Dick, D. M., Pulkkinen, L., Rose, R. J., & Kaprio, J. (2010). Tobacco, cannabis, and other illicit drug use among Finnish

adolescent twins: causal relationship or correlated liabilities?. *Journal of studies on alcohol and drugs*, *71*(1), 5.

Jaccard, J., Blanton, H., & Dodge, T. (2005). Peer influences on risk behavior: an analysis of the effects of a close friend. *Developmental Psychology, 41*(1), 135.

Jacobson, K. C., & Rowe, D. C. (1999). Genetic and environmental influences on the relationships between family connectedness, school connectedness, and adolescent depressed mood. *Developmental Psychology, 35*, 926–939.

Janssen, S. M., Chessa, A. G., & Murre, J. M. (2006). Memory for time: How people date events. *Memory & Cognition*, *34*(1), 138–147.

Jessor, R., & Jessor, S. L. (1977). *Problem behavior and psychosocial development: A longitudinal study of youth*. New York: Academic Press.

Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2013). Monitoring the Future national results on drug use: 2012 Overview, Key Findings on Adolescent Drug Use. *Ann Arbor: Institute for Social Research, The University of Michigan*.

Johnston, L.D., O'Malley, P.M., Bachman, J.G., Schulenberg, J.E., & Miech, R.A. (2014). Demographic subgroup trends among adolescents in the use of various licit and illicit drugs, 1975–2013 (Monitoring the Future Occasional Paper 81). *Ann Arbor, MI: Institute for Social Research.* Available: monitoringthefuture.org/pubs.html#papers

Jussim, L., & Osgood, D. W. (1989). Influence and similarity among friends: An integrative model applied to incarcerated adolescents. *Social Psychology Quarterly, 52*, 98–112.

Kendler, K. S., Schmitt, E., Aggen, S. H., & Prescott, C. A. (2008). Genetic and environmental influences on alcohol, caffeine, cannabis, and nicotine use from early adolescence to middle adulthood. *Archives of general psychiatry*, *65*(6), 674.

Klump, K. L., Culbert, K. M., Slane, J. D., Burt, S. A., Sisk, C. L., & Nigg, J. T. (2012). The effects of puberty on genetic risk for disordered eating: Evidence for a sex difference. *Psychological Medicine*, *42*, 627-637.

Koenis, M. M. G., Brouwer, R. M., van Baal, G. C. M., van Soelen, I. L. C., Peper, J. S., van Leeuwen, M., ... & Hulshoff Pol, H. E. (2013). Longitudinal study of hormonal and physical development in young twins. *The Journal of Clinical Endocrinology & Metabolism*, *98*(3), E518-E527.

Koopmans, J. R., Doornen, L. J., & Boomsma, D. I. (1997). Association between alcohol use and smoking in adolescent and young adult twins: A bivariate genetic analysis. *Alcoholism: Clinical and Experimental Research, 21*, 537-546.

Kretsch, N., Mendle, J., Cance, J. D., & Harden, K. P. (in press). Peer group similarity in perceptions of pubertal timing. *Journal of Youth and Adolescence.*

Kretsch, N., Mendle, J., & Harden, K. P. (in press). A twin study of objective and subjective pubertal timing and peer influence on risk-taking. *Journal of Research on Adolescence*.

Kretsch, N., Harden, K. P., Steinberg, L. D., & Tucker-Drob, E. M. (2014, March). A multidimensional model of adolescent risky decision-making: Integrating behavioral and self- report assessments. Poster presented at Society for Research on Adolescence biennial meeting, Austin, TX.

Kuhn, C., Johnson, M., Thomae, A., Luo, B., Simon, S.A., Zhou, G., & Walker, Q.D. (2010). The emergence of gonadal hormone influences on dopaminergic function during puberty. *Hormones and Behavior, 58,* 122–137.

Lahey, B. B. & D'Onofrio, B. M. (2010). All in the family: Comparing siblings to test causal hypotheses regarding environment influences on behavior. *Current Directions in Psychological Science, 19*, 319-323.

Lanza, S.T. & Collins, L.M. (2002). Pubertal timing and the onset of substance use in females during early adolescence. *Prevention Science, 3,* 69-81.

Lee, Y., & Styne, D. (2013). Influences on the onset and tempo of puberty in human beings and implications for adolescent psychological development. *Hormones and Behavior*, *64*(2), 250–261.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., ... & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, *8*, 75.

Lenz, B., Müller, C. P., Stoessel, C., Sperling, W., Biermann, T., Hillemacher, T., … & Kornhuber, J. (2012). Sex hormone activity in alcohol addiction: Integrating organizational and activational effects. *Progress in Neurobiology, 96,* 136-163.

Malina, R., & Bielicki, T. (1996). Retrospective longitudinal growth study of boys and girls active in sport. *Acta Paediatrica*, *85*(5), 570–576.

Marshall, W. A., & Tanner, J. M. (1968). Growth and physiological development during adolescence. *Annual Review of Medicine, 19*, 283–300.

Marshall, W. A., & Tanner, J. M. (1969). Variations in pattern of pubertal changes in girls. *Archives of Disease in Childhood, 44,* 291–303.

Marshall, W. A., & Tanner, J. M. (1970). Variations in the pattern of pubertal changes in boys. *Archives of Diseases in Childhood, 45,* 13–23.

Martin, C. A., Arch, G., Curry, T., & Martin, D. (1999). Alcohol use in adolescent females: correlates with estradiol and testosterone. The American Journal on Addictions, 8(1), 9-14.

Martin, C. A., Kelly, T. H., Rayens, M. K., Brogli, B. R., Brenzel, A., Smith, J., & Omar, H. A. (2002). Sensation seeking, puberty, and nicotine, alcohol and marijuana use in adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry, 41,* 1495-1502.

Matsueda, R. L., & Anderson, K. (1998). The dynamics of delinquent peers and delinquent behavior. *Criminology*, *36*(2), 269–308.

McClintock, M. K., & Herdt, G. (1996). Rethinking puberty: The development of sexual attraction. *Current Directions in Psychological Science, 5,* 178–183.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 415–444.

Mehta, P. H., Welker, K. M., Zilioli, S., & Carré, J. M. (2015). Testosterone and cortisol jointly modulate risk-taking. *Psychoneuroendocrinology*, *56*, 88-99.

Mehta, P. H., & Josephs, R. A. (2010). Testosterone and cortisol jointly regulate dominance: Evidence for a dual-hormone hypothesis. *Hormones and Behavior*, *58*(5), 898-906.

Mendle, J. (2014). Beyond Pubertal Timing New Directions for Studying Individual Differences in Development. *Current Directions in Psychological Science*, *23*(3), 215-219.

Mendle, J., Harden, K. P., Brooks-Gunn, J., & Graber, J. A. (2012). Peer relationships and depressive symptomatology in boys at puberty. *Developmental Psychology*, *48*(2), 429.

Mendle, J., Ferrero, J. (2012) Detrimental psychological outcomes associated with pubertal timing in adolescent boys. *Developmental Review*, *32*, 49-66.

Mendle, J., Turkheimer, E., & Emery, R.E. (2007). Detrimental psychological outcomes associated with early pubertal timing in adolescent girls. *Developmental Review, 27,* 151-171.

Michael, A., & Eccles, J. C. (2003). When coming of age means coming undone: Links between puberty and psychosocial adjustment among European American and African American girls. In C. Hayward (Ed.), Gender differences at puberty (pp. 277–303). New York: Cambridge University Press.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychological review*, *100*(4), 674.

Montoya, E. R., Bos, P. A., Terburg, D., Rosenberger, L. A., & van Honk, J. (2014). Cortisol administration induces global down-regulation of the brain's reward circuitry. *Psychoneuroendocrinology*, *47*, 31-42.

Moore, S.R., Harden, K.P., & Mendle, J. (2014). Pubertal timing and adolescent sexual behavior in girls. *Developmental Psychology, 50,* 1734-1745.

Muthén, L. K., & Muthén, B. O. (1998–2010). Mplus user's guide. (4th ed.). Los Angeles, CA: Muthen & Muthen.

Natsuaki, M. N., Biehl, M. C., & Ge, X. (2009). Trajectories of depressed mood from early adolescence to young adulthood: The effects of pubertal timing and adolescent dating. *Journal of Research on Adolescence*, *19*(1), 47–74.

Neale, M. C., & Cardon, L. R. (1992). Methodology for genetic studies of twins and families (NATO ASI Series D: Behavioral and Social Sciences, Vol. 67). Dordrecht, The Netherlands: Kluwer Academic.

Neighbors, C., Lee, C. M., Lewis, M. A., Fossos, N., & Larimer, M. E. (2007). Are social norms the best predictor of outcomes among heavy-drinking college students? *Journal of Studies on Alcohol, 67,* 290-299.

Nelson, E. E., Leibenluft, E., McClure, E. B., & Pine, D. S. (2005). The social reorientation of adolescence: A neuroscience perspective on the process and its relation to psychopathology. *Psychological Medicine, 35*, 163–174.

Nilsson, S., & Gustafsson, J. Å. (2000). Estrogen receptor transcription and transactivation: basic aspects of estrogen action. *Breast Cancer Research*, *2*, 360.

O'Brien, K. S., Caputi, P., Minto, R., Peoples, G., Hooper, C., Kell, S., & Sawley, E. (2009). Upward and downward physical appearance comparisons: Development of scales and examination of predictive qualities. *Body Image*, *6*(3), 201–206.

Ouimet, M. C., Simons-Morton, B. G., Zador, P. L., Lerner, N. D., Freedman, M., Duncan, G., Wang, J. (2010). Using the U.S. National Household Travel Survey to estimate the impact of passenger characteristics on young drivers' relative risk of fatal crash involvement. *Accident Analysis and Prevention, 42*, 689-694.

Ozer, E. M., Macdonald, T., & Irwin, C. E. J. (2002). Adolescent health care in the United States: Implications and projections for the new millennium. In J. T. Mortimer & R. W. Larson (Eds.), *Changing adolescent experience: Societal trends and the transition to adulthood.* Cambridge: Cambridge University Press.

Parent, A. S., Teilmann, G., Juul, A., Skakkebaek, N. E., Toppari, J., & Bourguignon, J. P. (2003). The timing of normal puberty and the age limits of sexual precocity: variations around the world, secular trends, and changes after migration. *Endocrine Reviews*, *24*(5), 668–693.

Peper, J. S., & Dahl, R. E. (2013). Surging hormones – Brain behavior interactions during puberty. *Current Directions in Psychological Science*, *22*, 134-139.

Peper, J. S., Koolschijn, C. M. P., & Crone, E. A. (2013). Development of risk taking: Contributions from adolescent testosterone and the orbito-frontal cortex. *Journal of Cognitive Neuroscience.*

Peters, S., Jolles, D. J., Van Duijvenvoorde, A. C., Crone, E. A., & Peper, J. S. (2015). The link between testosterone and amygdala-orbitofrontal cortex connectivity in adolescent alcohol use. *Psychoneuroendocrinology*.

Petersen, A. C., Crockett, L., Richards, M., & others. (1988). A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence*, *17*(2), 117–133.

Petersen, A. C., & Taylor, B. (1980). The biological approach to adolescence: Biological change and psychological adaptation. In J. Adelson (Ed.), Handbook of adolescent psychology. New York: Wiley.

Pickles, A., Neale, M., Simonoff, E., Rutter, M., Meyer, J., Crouchley, R., Silberg, J., & Eaves, L. (1994). A simple method for censored age-of-onset data subject to recall bias. *Behavior Genetics, 24,* 457-468.

Popma, A., Vermeiren, R., Geluk, C. A., Rinne, T., van den Brink, W., Knol, D. L., ... & Doreleijers, T. A. (2007). Cortisol moderates the relationship between testosterone and aggression in delinquent male adolescents. *Biological Psychiatry*, *61*(3), 405-411.

Primus, R. J., & Kellogg, C. K. (1989). Pubertal-related changes influence the development of environment-related social interaction in the male rat. *Developmental Psychobiology, 22(6),* 633–643.

Resnick, M. D., Bearman, P. S., Blum, R. W., Bauman, K. E., Harris, K. M., Jones, J., . . . Udry, R. (1997). Protecting adolescents from harm: Findings from the National Longitudinal Study on Adolescent Health. *Journal of the American Medical Association, 278,* 823–832.

Reyna VF, Farley F. (2006). Risk and rationality in adolescent decision-making: Implications for theory, practice, and public policy. *Psychological Science in the Public Interest, 7,* 1–44.

Reynolds, B. M., & Juvonen, J. (2011). The role of early maturation, perceived popularity, and rumors in the emergence of internalizing symptoms among adolescent girls. *Journal of Youth and Adolescence*, *40*(11), 1407-1422.

Reynolds, M. D., Tarter, R., Kirisci, L., Kirillova, G., Brown, S., Clark, D. B., & Gavaler, J. (2007). Testosterone levels and sexual maturation predict substance use disorders in adolescent boys: A prospective study. *Biological Psychiatry, 61*, 1223-1227.

Rierdan, J., Koff, E., & Stubbs, M. L. (1988). Timing of menarche, preparation, and initial menstrual experience: Replication and further analyses in a prospective study. *Journal of Youth and Adolescence, 18,* 413– 426.

Rose, A. J., & Rudolph, K. D. (2006). A review of sex differences in peer relationship processes: potential trade-offs for the emotional and behavioral development of girls and boys. *Psychological Bulletin*, *132*(1), 98.

Rose, R. J., & Dick, D. M. (2004). Gene-environment interplay in adolescent drinking behavior. *Alcohol Research and Health*, *28*(4), 222.

Rowe, R., Maughan, B., Worthman, C. M., Costello, E. J., & Angold, A. (2004). Testosterone, antisocial behavior, and social dominance in boys: Pubertal development and biosocial interaction. *Biological Psychiatry*, *55*, 546–552.

Sapienza, P., Zingales, L., & Maestripieri, D. (2009). Gender differences in financial risk aversion and career choices are affected by testosterone. *Proceedings of the National Academy of Sciences*, *106*(36), 15268-15273.

Schulz, K. M., & Sisk, C. L. (2006). Pubertal hormones, the adolescent brain, and the maturation of social behaviors: Lessons from the Syrian hamster. *Molecular and Cellular Endocrinology, 254–255,* 120–126.

Schulz, K.M., Molenda-Figueira, H.A., Sisk, C.L., 2009. Back to the future: the organizational-activational hypothesis adapted to puberty and adolescence. *Hormones and Behavior 55*, 597–604.

Shelton, K. H., & Van Den Bree, M. (2010). The moderating effects of pubertal timing on the longitudinal associations between parent–child relationship quality and adolescent substance use. *Journal of Research on Adolescence*, *20*(4), 1044-1064.

Shirtcliff, E. A., Dahl, R. E., & Pollak, S. D. (2009). Pubertal development: Correspondence between hormonal and physical development. *Child Development, 80*, 327–337.

Sickmund, M., & Puzzanchera, C. (2014). Juvenile offenders and victims: 2014 National report.

Silberg, J. L., Parr, T., Neale, M. C., Rutter, M., Angold, A., & Eaves, L. J. (2003). Maternal smoking during pregnancy and risk to boys' conduct disturbance: an examination of the causal hypothesis. *Biological Psychiatry, 53*(2), 130-135.

Sisk, C. L., & Zehr, J. L. (2005). Pubertal hormones organize the adolescent brain and behavior. *Frontiers in Neuroendocrinology, 26(3–4),* 163–174.

Sjöberg, R. L., Ducci, F., Barr, C. S., Newman, T. K., Dell'Osso, L., Virkkunen, M., & Goldman, D. (2007). A non-additive interaction of a functional MAO-A VNTR and testosterone predicts antisocial behavior. *Neuropsychopharmacology*, *33*, 425-430.

Skoog, T., Sorbring, E., & Bohlin, M. (2015). Facebook as a means to make new peers among early maturing girls. *Computers in human behavior*, *48*, 500-505.

Slater, M. D., Kelly, K. J., Lawrence, F. R., Stanley, L. R., & Comello, M. L. G. (2011). Assessing media campaigns linking marijuana non-use with autonomy and aspirations:"Be Under Your Own Influence" and ONDCP's "Above the Influence". *Prevention Science, 12*(1), 12-22.

Snijders, T. A. B., Steglich, C. E. G., & Van de Bunt, G. G. (2010). Introduction to actor-based models for network dynamics. *Social Networks, 32,* 44-60.

Somerville, L. H., Jones, R. M., & Casey, B. J. (2010). A time of change: Behavioral and neural correlates of adolescent sensitivity to appetitive and aversive environmental cues. *Brain and Cognition, 72*, 124-133.

Spear, L. P. (2000). The adolescent brain and age-related behavioral manifestations. *Neuroscience and Biobehavioral Reviews, 24*, 417–463.

Spear, L. (2013). The Teenage Brain Adolescents and Alcohol. *Current Directions in Psychological Science*, *22*(2), 152-157.

Spitz, E., Moutier, R., Reed, T., Busnel, M. C., Marchaland, C., Roubertoux, P. L., & Carlier, M. (1996). Comparative diagnoses of twin zygosity by SSLP variant analysis, questionnaire, and dermatoglyphic analysis. *Behavior Genetics, 26*, 55–63.

Stanton, S. J., Liening, S. H., & Schultheiss, O. C. (2011). Testosterone is positively associated with risk taking in the Iowa Gambling Task. *Hormones and Behavior, 59*, 252-256.

Stattin, H., Kerr, M., & Skoog, T. (2011). Early pubertal timing and girls' problem behavior: Integrating 2 hypotheses. *Journal of Youth and Adolescence, 40,* 1271-1287.

Stattin, H., & Magnusson, D. (1990). *Pubertal maturation in female development.* Hillsdale, NJ: Erlbaum.

Steinberg, L. (2007). Risk taking in adolescence: New perspectives from brain and behavioral science. *Current Directions in Psychological Science, 16,* 55–59.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review, 28*, 78-106.

Steinberg, L., & Monahan, K. (2007). Age differences in resistance to peer influence. *Developmental Psychology, 43,* 1531-1543.

Stice, E., Presnell, K., & Bearman, S.K. (2001). Relation of early menarche to depression, eating disorders, substance abuse, and comorbid psychopathology among adolescent girls. *Developmental Psychology, 37,* 608-619.

Strasburger, V. C., & Brown, R. T. (1991). *Adolescent medicine: A practical guide*. Boston, MA: Little, Brown.

Substance Abuse and Mental Health Services Administration, *Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings*, NSDUH

Series H-48, HHS Publication No. (SMA) 14-4863. Rockville, MD: Substance Abuse and Mental Health Services Administration, 2014.

Sun, S. S., Schubert, C. M., Chumlea, W. C., Roche, A. F., Kulin, H. E., Lee, P. A., … Ryan, A. S. (2002). National estimates of the timing of sexual maturation and racial differences among US children. *Pediatrics*, *110*(5), 911–919.

Susman, E. J., & Rogol, A. (2004). Puberty and psychological develop- ment. In R. M. Lerner, & L. Steinberg (Eds.), *Handbook of adolescent psychology* (2nd ed., pp. 15–44). Hoboken, NJ: Wiley.

Tackett, J. L., Herzhoff, K., Harden, K. P., Page-Gould, E., & Josephs, R. A. (2014). Personality× hormone interactions in adolescent externalizing psychopathology. *Personality Disorders: Theory, Research, and Treatment*, *5*(3), 235.

Tarter, R. E., Kirisci, L., Gavaler, J. S., Reynolds, M., Kirillova, G., Clark, D. B., ... & Vanyukov, M. (2009). Prospective study of the association between abandoned dwellings and testosterone level on the development of behaviors leading to cannabis use disorder in boys. *Biological Psychiatry*, *65*(2), 116-121.

Tolson, J. M., & Urberg, K. A. (1993). Similarity between adolescent best friends. *Journal of Adolescent Research*, *8*(3), 274–288.

Udry, J. R. (1991). Predicting alcohol use by adolescent males. *Journal of Biosocial Science, 23,* 381–386.

Udry, J. R. (2003). The National Longitudinal Study of Adolescent Health (Add Health), Waves I & II, 1994–1996; Wave III, 2001– 2002 [machine-readable data file and

documentation]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill.

Urberg, K. A., Değirmencioğlu, S. M., Tolson, J. M., & Halliday-Scher, K. (1995). The structure of adolescent peer networks. *Developmental Psychology*, *31*(4), 540.

van Honk, J., Schutter, D. J., Hermans, E. J., Putman, P., Tuiten, A., & Koppeschaar, H. (2004). Testosterone shifts the balance between sensitivity for punishment and reward in healthy young women. *Psychoneuroendocrinology, 29*, 937-943.

Van Den Bos, W., Golka, P. J., Effelsberg, D., & McClure, S. M. (2013). Pyrrhic victories: the need for social status drives costly competitive behavior. *Frontiers in Neuroscience*, *7*.

Vermeersch, H., T'sjoen, G., Kaufman, J. M., & Vincke, J. (2008). Estradiol, testosterone, differential association and aggressive and non-aggressive risk-taking in adolescent girls. *Psychoneuroendocrinology, 33,* 897-908.

Vetter‐O'Hagen, C. S., & Spear, L. P. (2012). Hormonal and physical markers of puberty and their relationship to adolescent‐typical novelty‐directed behavior. *Developmental Psychobiology*, *54*(5), 523-535.

Viau, V. (2002). Functional cross‐talk between the hypothalamic‐pituitary‐gonadal and‐adrenal axes. *Journal of Neuroendocrinology*, *14*(6), 506-513.

Wichstrom, L. (2001). The impact of pubertal timing on adolescents' alcohol use. *Journal of Research on Adolescence, 11*, 131–150.

Wilson, D.M., Killen, J.D., Hayward, C., Robinson, T.N., Hammer, L.D., Kraemer, H.C., Varady, A., Taylor, C.B., 1994. Timing and rate of sexual maturation and the onset of cigarette and alcohol use among teenage girls. Arch. Pediatr. Adolesc. Med. 148, 789–795.

Witt, E. D. (2007). Puberty, hormones, and sex differences in alcohol abuse and dependence. *Neurotoxicology and Teratology, 29*, 81-95.

Wood, R.I., 2004. Reinforcing aspects of androgens. *Physiology and Behavior, 83 (2)*, 279–289.

Young, S. E., Friedman, N.P., Miyake, A., Willcutt, E.G., Corley, R. P., Haberstick, B.C., & Hewitt, J. K. (2000). Genetic and environmental influences on behavioral disinhibition. *Journal of Abnormal Psychology, 118*, 117-130.

Yuan, A. S. V. (2007). Gender differences in the relationship of puberty with adolescents' depressive symptoms: Do body perceptions matter? *Sex Roles*, *57*(1-2), 69–80.

Yun, I., Cheong, J., & Walsh, A. (2011). Genetic and environmental influences in delinquent peer affiliation: From the peer network approach. *Youth Violence and Juvenile Justice, 9,* 241-258.

Zimring, F. E. (1998). American youth violence. Oxford, England: Oxford University Press.

Zuckerman, M., Buchsbaum, M. S., & Murphy, D. L. (1980). Sensation seeking and its biological correlates. *Psychological Bulletin, 88*, 187-214.